Numerical Weather Prediction

An assessment of the impact of new physical parameterisations on the performance of the Global Unified Model in THORPEX 15-day forecasts.

Forecasting Research Technical Report No. 496

Nicholas Savage, Sean Milton, David Walters and Julian Heming

email:nwp_publications@metoffice.gov.uk

© Crown copyright

An assessment of the impact of new physical parameterisations on the performance of the Global Unified Model in THORPEX 15-day forecasts

Nicholas Savage, Sean Milton, David Walters and Julian Heming

February 2006

Abstract

The impact of a package of improved physical parameterisations, primarily in the marine boundary layer and the convection scheme on the global Unified Model performance at forecast ranges up to 15 days have been evaluated. In general model performance is improved with reduced systematic errors in many regions. In particular, large improvements in the model in the winter northern hemisphere surface cold bias and to the tropical divergent circulation were seen. The Madden-Julian Oscillation in the revised model is also improved. However the warm bias in summer over North America and Asia is made worse and precipitation and temperature anomalies over Africa and South America increase. A comparison of the revised model version to ECMWF model results has also been carried out and although in general the systematic errors in the Unified Model results are slightly larger, the comparison is overall encouraging for multi-model ensemble forecasting.

1 Introduction

THe **O**bserving System **R**esearch and **P**redictability **EX**periment (THORPEX) is a ten-year long international research programme to accelerate improvements in the accuracy of 1 to 14-day weather forecasts for the benefit of society, the economy and the environment. It is being developed and implemented as a part of the WMO[1] World Weather Research Programme (WWRP). For more details see Shapiro and Thorpe (2004). The Met Office is currently running a project to evaluating the benefits of engaging with THORPEX and is running in near-real time a fifteen day ensemble prediction system. The forecast model is based on the global component of the Met Office Global and Regional Ensemble Prediction System (MOGREPS, Bowler et al, 2007a,b), and a 24-member ensemble is run twice daily (00Z and 12Z) out to 15 days using part of the UK allocation on the ECMWF supercomputer. The initial conditions and perturbation fields are obtained from MOGREPS (Global). The objective of this work is to investigate possible improvements in the performance of the 15 day forecasts by implementing recent changes made to the physics of the operational global version of the Unified Model (UM).

A preliminary assessment of the performance of the Unified Model at 15 day timescales has already been carried out (Milton *et al*, 2007). This report builds on the earlier work to evaluate the impact of including new physical parameterisations on the model performance at these timescales. A package of improvements to model physics in the global operational version of the UM were introduced on the 14th March 2006. The principal aims of the package were to improve the tropical performance of the UM and in particular to reduce errors in:

- sea surface evaporation;
- precipitation over the sea;
- BL winds over the sea; and
- the vertical profiles of temperature and humidity.

[1] World Meteorological Organisation

Winter	Summer
26.11.2003	14.06.2003
14.12.2003	03.07.2003
03.01.2004	22.07.2003
22.01.2004	11.08.2003
11.02.2004	30.08.2003
21.11.2005	03.06.2006
15.12.2005	23.06.2006
04.01.2006	11.07.2006
24.01.2006	02.08.2006
13.02.2006	18.08.2006

Table 1: Initial analysis dates (12 UTC) for the case study forecasts

For more details of the changes to model physics and their impact on the short range forecasts see Willet (2006).

Section 2 describes in more detail the model formulations and the test cases which have been run, section 3 presents the results with a focus on the change in model performance when using the new physics parameterisations, section 4 compares system errors of the UM and ECMWF models, and finally section 5 draws conclusions about whether the new physics package should be implemented in the 15 day ensemble forecasts and discusses further work.

2 Model formulation and test cases

All experiments described in this report were run at a horizontal resolution of N144[2] (288 longitude points, 217 latitude points) which is the same as the resolution of the MOGREPS-15 ensemble forecast system run at ECMWF. The model contains 38 levels in the vertical with the model top at 38 km. The current operational version of the model is N320 with 50 levels in the vertical and a model top at 65km.

2.1 Test cases

Twenty N144 15-day global model forecasts initialised from Met Office 12UTC operational analyses. The cases, ten summer and ten winter, have been run with both model formulations and details of the cases are shown in Table 1. The initial analyses dates are at least 16 days apart so that the cases do not overlap and have some synoptic independence. The initial analyses in 2003/4 were produced with a 3D variational assimilation system (Lorenc *et al*, 2000) while those for 2005/6 were produced with a 4D variational analysis scheme.

For forecasts of beyond a week the values of boundary conditions such as sea surface temperature (SST), soil moisture content (SMC) and sea ice fields are of increasing importance for predictability. The initial fields of these variables differ between the two runs.

2.2 HadGEMminus (G34-) model formulation

The "control" model formulation against which the version with the new physics package is compared in this work differs somewhat from that used in Milton *et al*, 2007. The "control" model here is based more closely on that currently run in near real time at ECMWF. It is very similar to model cycle G34 as described in Milton *et al* (2005) (introduced in the operational model 18/1/2005 and with no further physics changes until cycle G39 14/3/2006). This version also forms the major components of the coupled climate configuration HadGEM1 (Martin et al, 2005). However, for technical reasons associated with the stochastic physics scheme the model configuration used here is run without the restructured convection scheme. It differs from the operational model in that it had a fixed CAPE closure timescale of 30 minutes instead of a timescale which was reduced at high relative humidities (see comments later). Finally this version does not use the increased Saharan albedos introduced at model cycle G34. This version is therefore known as 'HadGEMminus'. All ancillary files (ozone, snow depth, soil moistures, deep soil temperature,

[2]This equates to half the number of points around a latitude circle or to the number of resolved (2 grid length) waves

soil parameters, vegetation distribution parameters, sea surface temperatures, and sea ice fields) were taken from climatological values to reproduce the version of the THORPEX forecasts (this has since changed to using value from the operational model analysis). The details of climatologies given here are based on Jones (2004). Both the sea surface temperature and sea ice concentration fields have been derived from GISST 2.0 (Global sea-Ice and Sea Surface Temperature) climatology (Parker *et al* 1995). Each has 12 fields valid at the middle of the month. For each sea ice concentration field there are also derived fields of fractional time and sea ice thickness. Soil moisture and snow amount is derived from the Willmott *et al* (1985) climatology. The deep soil temperature climatology used is that created by a 17 year AMIP run of the climate model. Soil and vegetation parameters are based on the soils classification dataset created by Wilson and Henderson-Sellers (1985) (WHS) and the ozone climatology is based on Li and Shine (1995). Sea ice and sea surface temperatures plus ozone fields are fixed throughout each run both in this model formulation and in the runs of the revised model formulation, while other fields such as vegetation and soil moisture are allowed to evolve. At the time this experiment was started initial values of snow depth; soil moisture; deep soil temperature; sea surface temperatures; sea ice fraction and sea ice thickness for the ensemble prediction system were based on climatologies but have since been changed to use fields reconfigured from the global forecast model.

The model configuration used for the initial THORPEX tests in Milton et al (2007) had a CAPE closure timescale with a maximum of 1 hour with CAPE timescale reduced at high relative humidities, used the older 8A boundary layer and 3B microphysics and had some differences from this version in the dynamics settings. By using this new configuration (HadGEMminus) as our control we can better evaluate the effect of the physics package on the ensemble prediction system being run at ECMWF. The only differences with the version used at ECMWF are very small impacts from the Portable Unified Model (PUM) mods and that the MOGREPS start dumps used for the THORPEX system use reconfigured snow depth and sea ice thickness from the operational model.

2.3 PS11 (G39) model formulation

The second global model formulation used in these experiments was that implemented in the Met Office operational global forecast model in spring 2006 (model cycle G39). The major changes to the physics introduced with this set of updates to the global model were focused on improving tropical model performance. This discussion of the model changes draws heavily on Willet (2006).

The changes were to the boundary layer scheme and convection. The changes to the convection scheme were:

- some small changes for improved numerical performance to mid level convection which had little impact on model results and
- the introduction of an adaptive detrainment scheme. The triggering of forced detrainment is set to be a function of the parcel's buoyancy gradient resulting in a much smoother detrainment and mass flux profile (Maidens and Derbyshire, 2007).

The boundary layer changes were aimed at the marine BL and were:

- Sharp tails over sea. A rapid decline in amount of turbulence is observed as the stability of the atmosphere increases. This stability dependence is referred to as the 'tail' of the mixing function: a rapid decline in the amount of turbulence with stability is referred to as a 'sharp tail' and a slow decline as a 'long tail'. Over land it is necessary to use unrealistic 'long tails' because of the variability of the land surface. Over the sea, which has a consistent surface type, the model now uses 'sharp tails'. The main benefits are an increased turning of the wind in the boundary layer and a reduction in near-surface wind speed in stable boundary layers over the sea (Brown et al, 2007).;
- The formulation of the thermal roughness length has been changed, so that at high wind speeds the surface exchange coefficient is in better agreement with observational data. The slower rate of evaporation from salty water is also now taken into account. These two changes reduce the evaporation from the sea surface and dry the near surface atmosphere, and hence improve the model's agreement with observational data. For more information see Edwards (2007).
- revision to the diagnosis of decoupled boundary layers;

- an improvement to the diagnosis of cumulus. The existing, thermodynamic, diagnosis of a cumulus-capped boundary layer (and so the triggering of the convection scheme from the lifting condensation level) was prevented for near-neutral marine boundary layers, i.e.. regimes of strong wind shear over the sea.
- correction to prevent entrainment in stable BLs and
- inclusion of non-gradient stress. Transport by turbulent eddies which is not represented by the gradient flux has been extended from the temperature flux to include the moment flux. See Brown and Grant (1997) and Brown et al (2007).

In addition to these code changes the following ancillary fields were reconfigured from the operational model analyses and therefore should provide a more realistic initial estimate of boundary forcing: snow depth; soil moisture; deep soil temperature; sea surface temperatures; sea ice fraction and sea ice thickness. The remaining ancillary files which used climatologies were: ozone; soil parameters and vegetation distribution parameters. The soil parameter ancillary files were updated to those used in the operational model since cycle G34 which is a modified version of the WHS data but increases the bare soil albedo over the Sahara to give more realistic values. Otherwise the ancillaries are essentially unchanged from those used for the HadGEMminus runs. See Milton *et al* (2005) for more details of the revised albedos. This version also had a default CAPE closure timescale of 30 minutes with a minimum equal to the convection time-step and a relative humidity threshold of 60% below which the CAPE closure timescale is reduced plus the restructured convection scheme used for HadGEM1.

3 Results

This section contains the bulk of the results and consists mostly of a comparison of model performance between the two model configurations with the main objective of reaching a conclusion as to whether the new model configuration should be implemented for the 15 day forecast system.

3.1 Temperature

Northern hemisphere winter

Figure 1 shows the growth with time of the zonal average mean errors at 1000 and 250 hPa. The most dramatic difference between the two configurations in the 1000hPa fields in the Northern extra-tropics. In the HadGEMminus model configuration a large cold bias of up to -8K develops by day 5 between 60 and 90 north. In the PS11 runs this is much reduced and in some parts of this area there is a small warm bias (maximum of approx. +1K). Much of the area in which the temperature is higher than in the control corresponds to a region where total cloud amounts increase by more than 0.1 (see below). Therefore reduced radiative cooling at the surface in the PS11 runs is a possible explanation of the warmer near surface temperatures in this region. The errors at other latitudes at this level are similar between the two configurations with errors in both versions typically between -1 and +1 K.

At 250 hPa the changes in temperature error are more prominent in the tropics with the cold bias being reduced from around -1K to near zero. At high northern latitudes there is also a warming of up to 2K in the PS11 model version compared to the HadGEMminus. This warming seems to be making the model agreement with the analysis worse but it should be remembered that this is a fairly small area.

These results can be also seen in the zonal mean difference plots from the mean values over days 11-15 (figure 2). In general the PS11 model version gives much better temperatures in the northern mid-latitudes with the strong near surface cold bias being removed and the temperature structure also being improved. These large changes in surface temperature are consistent with changes in cloudiness between the two runs (see section 3.8)

Both models have strong positive temperature biases (>+3K) in the tropical stratosphere which are very similar. This is consistent with the tentative conclusion in Milton *et al* (2007) that this was a result of the ozone climatology used which is identical in the two model configurations. Further tests with alternative ozone climatologies are planned.

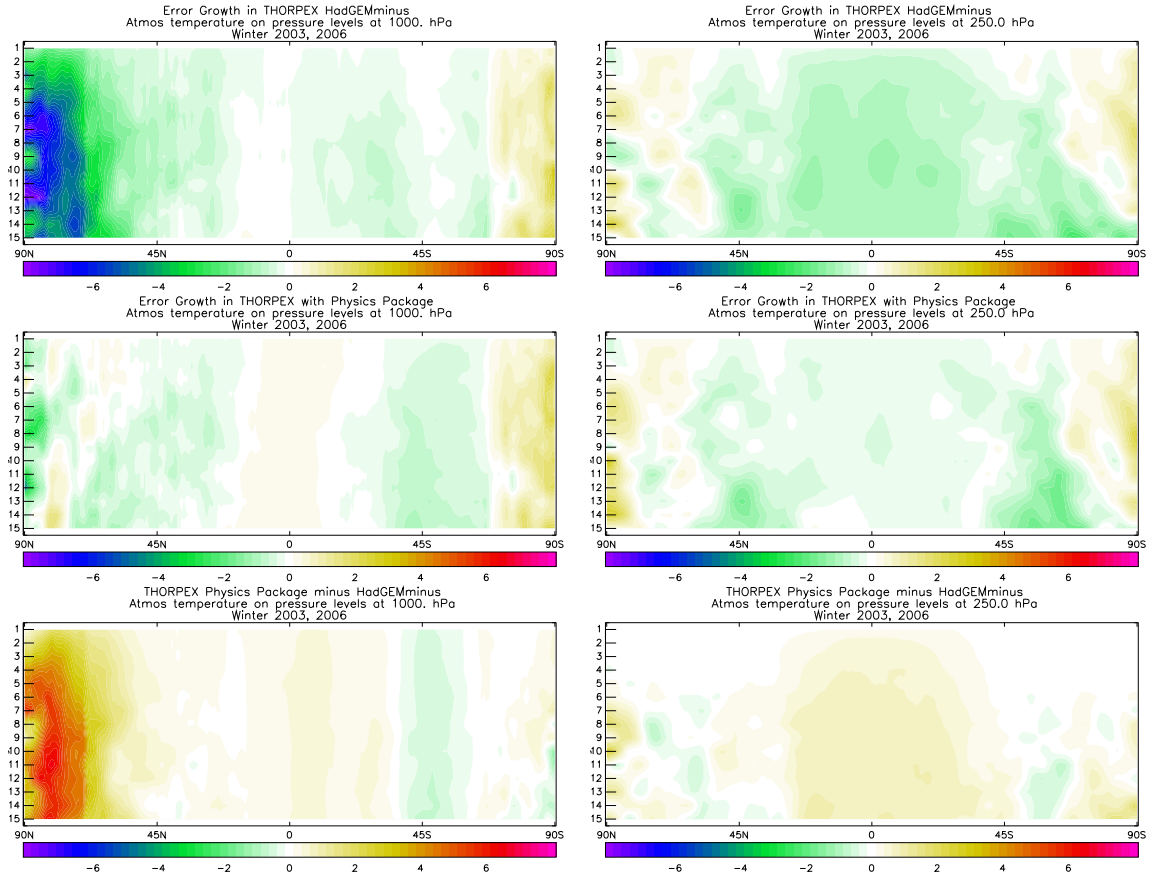

Figure 1: Error growth in zonal mean temperature field at 1000hPa (left) and 250hPa

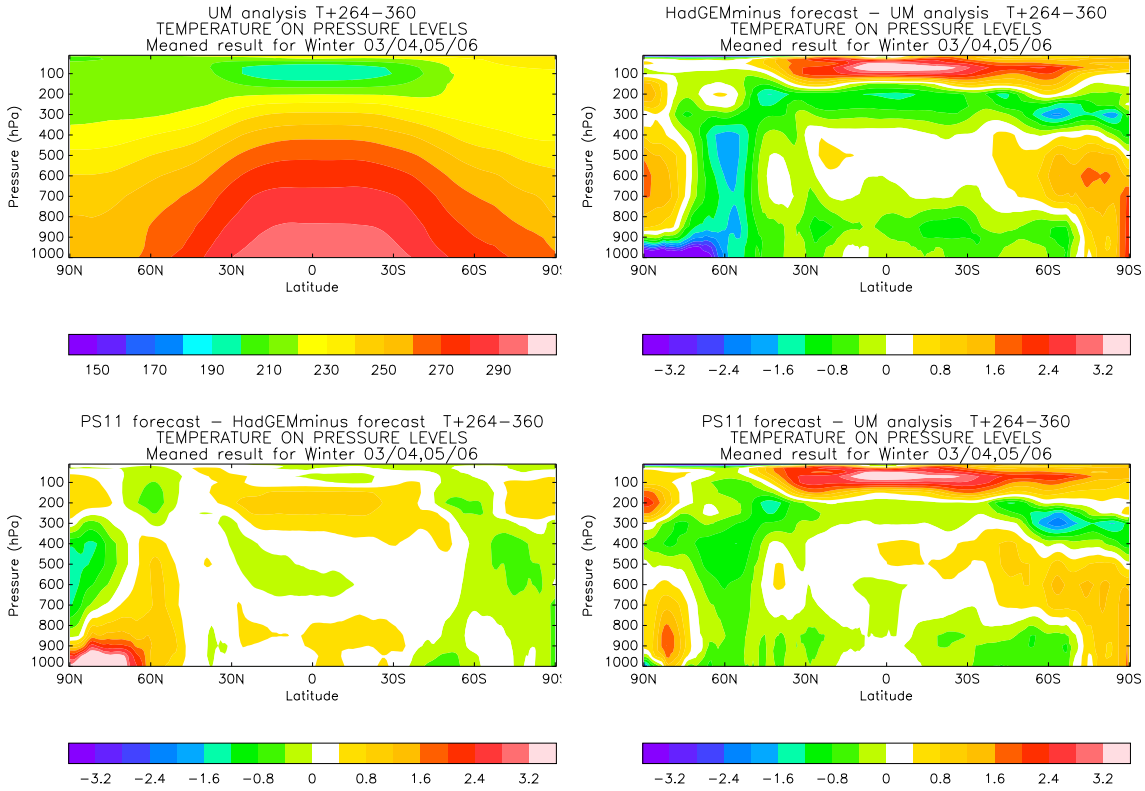

Figure 2: Zonal mean bias averaged over days 11-15 of all winter cases

Northern hemisphere summer

Figure 3 shows the zonal mean temperature biases in the June to August season for days 11-15 in both models. The differences between the two models in the tropics are very similar to winter season. There is again a decrease in the cold bias at high northern latitudes but it substantially smaller than in the winter. In contrast there appears to be an increase in the cold bias at about 400 hPa above Antarctica in PS11. The new model formulation also shows a warm bias near the surface between 30N and 60N which is not present in the control. This is a result of the worsening of the surface warm bias over Asia and North America (see later). Figure 4 shows the verification of both models versus sonde data at T+360. The reduction of the cold surface bias is very clear in the northern hemisphere with corresponding large decreases in RMS error at most levels and most prominently at the surface. In the tropics the improvement is not as obvious with PS11 actually having higher RMS errors at lower levels but generally giving a lower bias at most levels than HadGEMminus. At T+120 (not shown) the changes in tropical temperatures are consistent with those found in Willet (2006) with the mid-tropospheric (600 hPa-400 hPa) warm bias and the upper tropospheric (250 hPa-200 hPa) cold bias much reduced, and large reductions in root-mean-square (r.m.s.) temperature error throughout the mid and upper troposphere. At T+360 the two model versions have the same warm bias at 500 hPa but this is reached more slowly. This suggests that different mechanisms may be responsible for this bias in the two model formulations.

The temperature distribution at the surface (figure 5) indicates some additional features. There is a reduction in the cold bias over Central and Western Europe, which may be due to drier soil moistures in the over Europe in summer 2006 caused by soil moisture nudging in the UM. There are also smaller cold biases in Northern Russia, Australia and parts of Africa with typical increases in temperature in these regions of ~2K. However the warm bias in Asia and North America is made worse with positive temperature anomalies of more than 3 K. This signal is also seen in the operational model during summer 2006 and is under investigation.

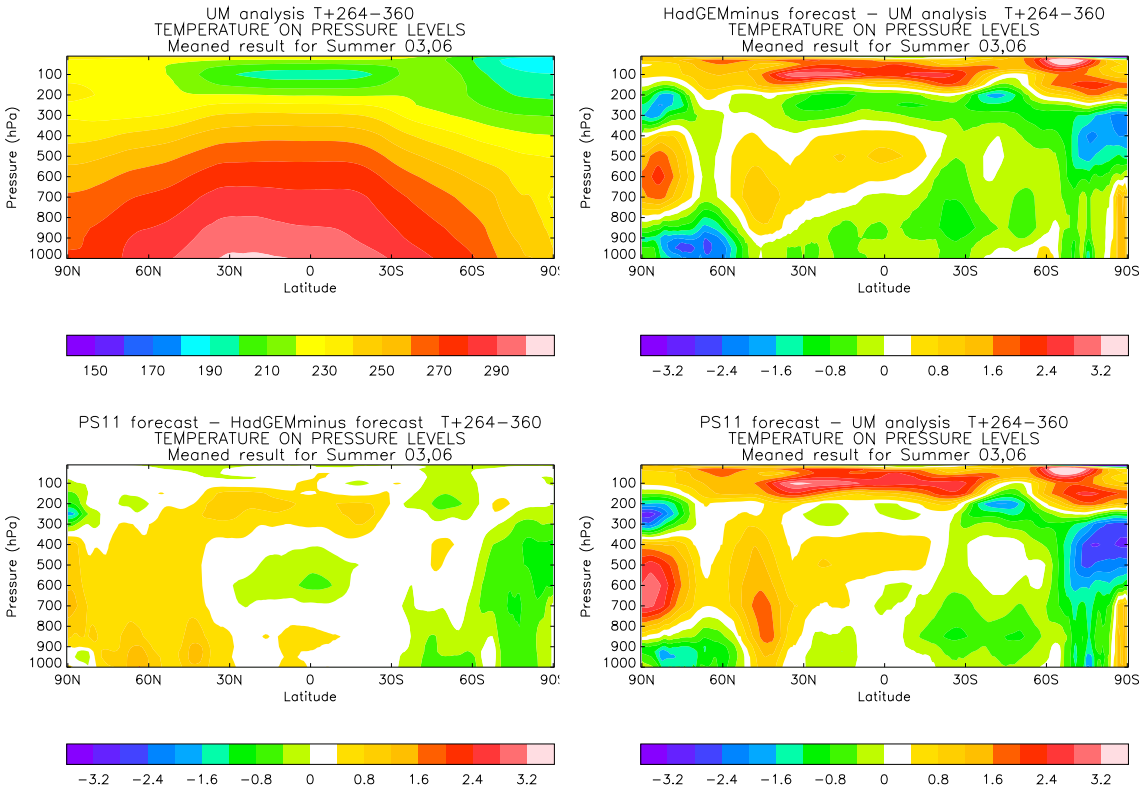

Figure 3: As Figure 2 except for Summer Cases

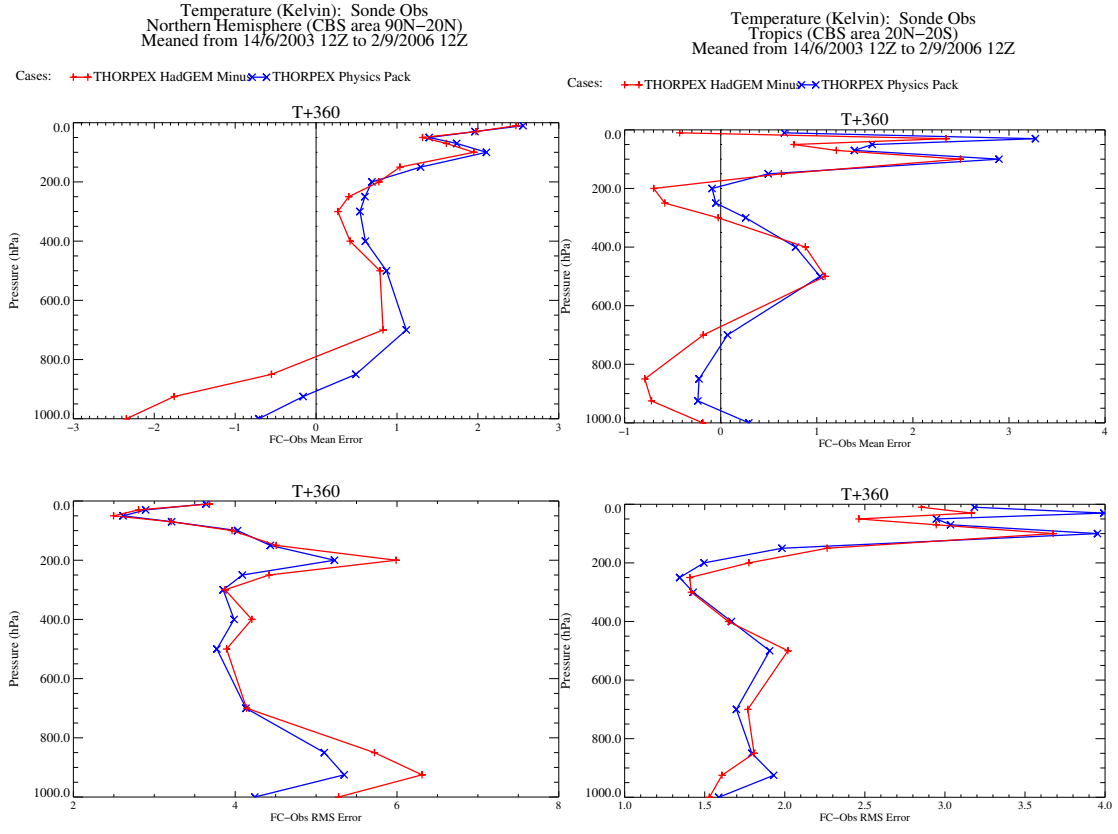

Figure 4: Verification of temperature profiles against radiosonde data in the northern hemisphere (left) and tropics (right) at day 15.

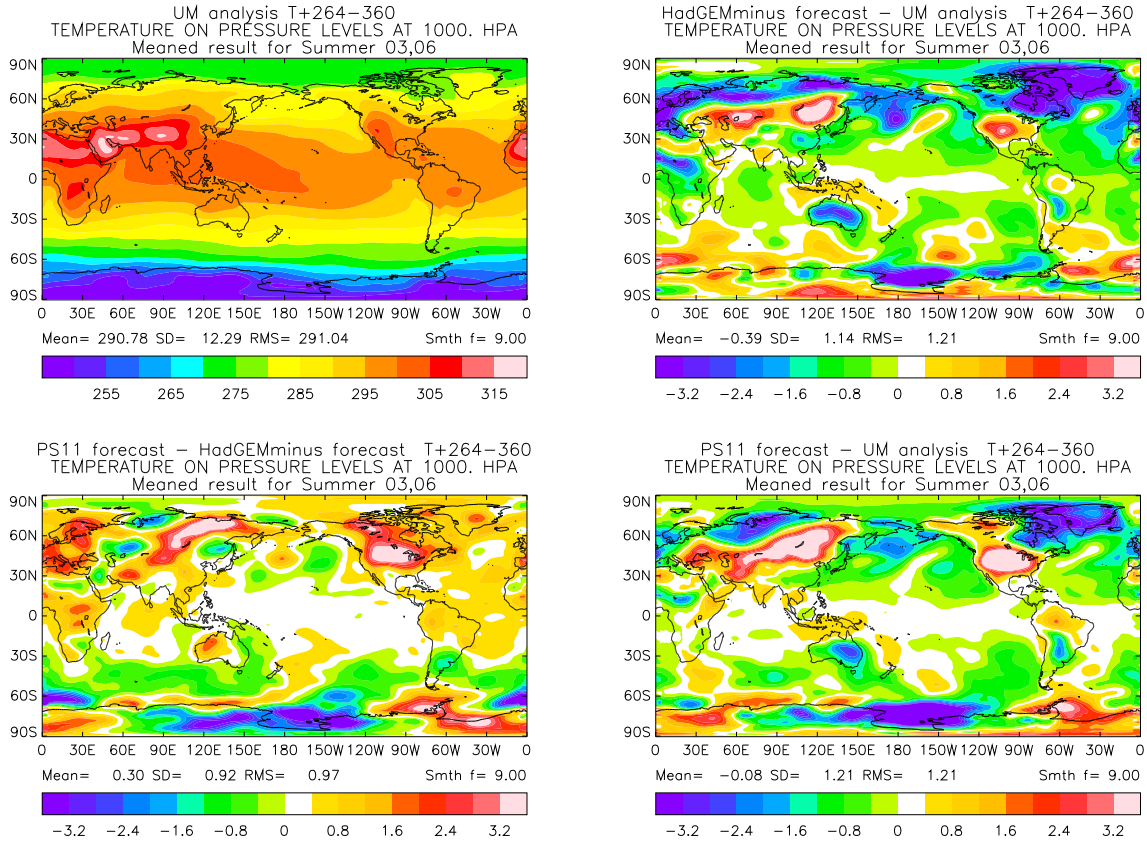

Figure 5: Temperature biases at 1000hPa averaged over days 11-15 of all JJA cases.

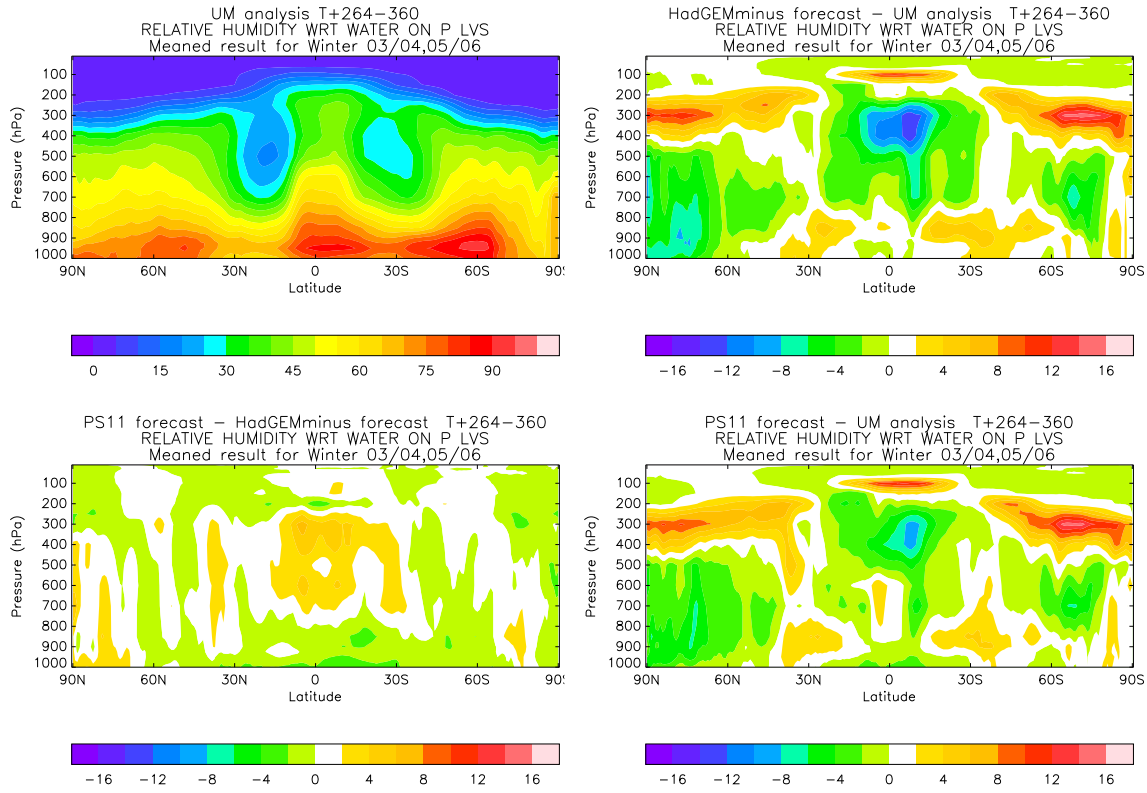

Figure 6: Zonal mean relative humidity biases averaged over days 11-15 of all Winter cases

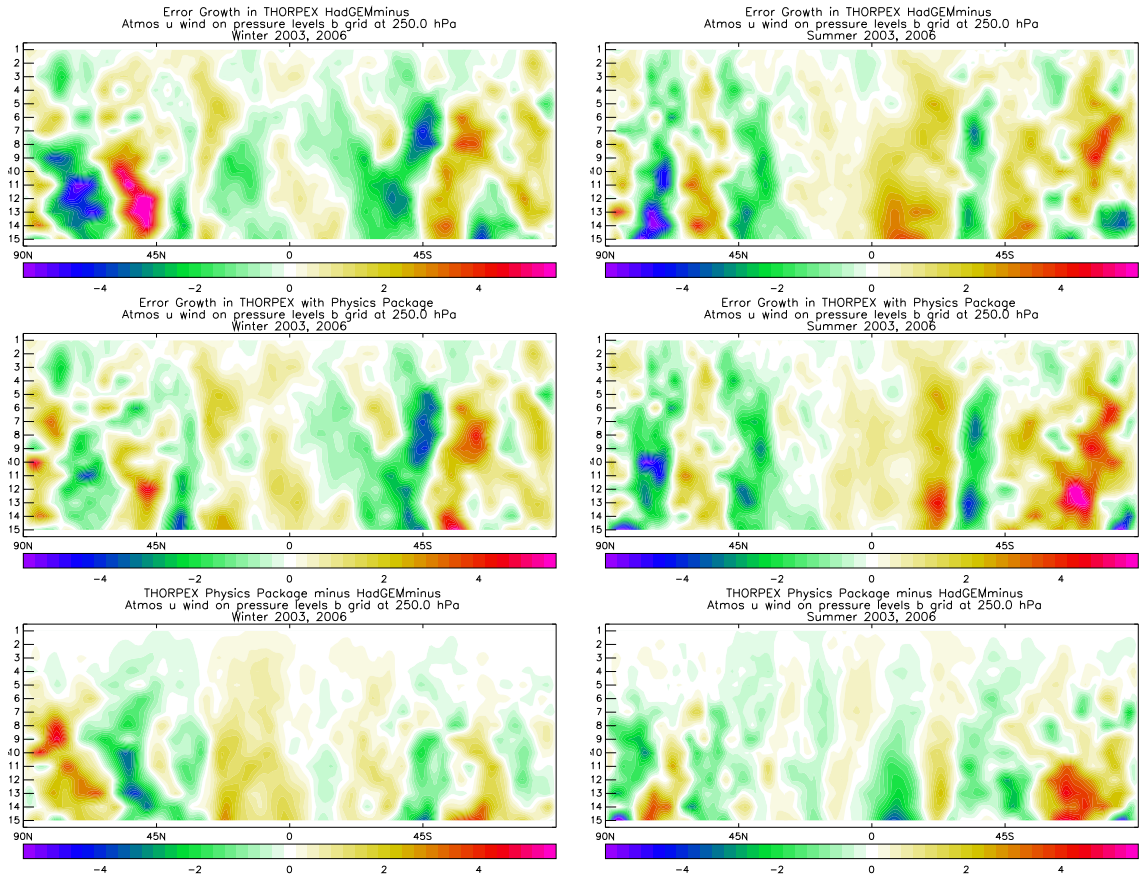

Figure 7: Zonal mean error growth plots for zonal mean u wind at 250 hPa for DJF (left) and JJA (right)

3.2 Relative humidity

Figure 6 shows the zonal mean bias in relative humidity for both model configurations averaged over days 11 to 15 of all the winter cases. There is a reduction in the upper tropical troposphere dry bias from a peak of over -12% to around -8% consistent with the expected effects of the inclusion of adaptive detrainment in the convective scheme. Elsewhere the changes from HadGEMminus to PS11 are generally positive. However the large (>10%) positive bias in relative humidity near the tropopause is still present and large unchanged by these changes.

3.3 Zonal and meridional winds

Zonal winds

Figure 7 shows the error growth in 250 hPa zonal mean zonal wind speed in winter and summer and figure 8 shows the zonal mean wind biases in DJF. The error growth to about seven days is similar between the two model configurations with differences of less than 0.5 ms^{-1}. However after 8 days the differences between the configurations rapidly increase in both seasons. In winter there is a large decrease in the westerly bias seen in HadGEMminus at 50 N by about -4 $ms^{-1}$and the easterly bias further north is also reduced in the PS11 model version. In the northern hemisphere summer there is a similar reduction of the northern hemisphere errors in middle and high latitudes in the PS11 version with peak biases in zonal wind speeds falling from -5 to +4 ms^{-1} to -4 to +2 ms^{-1}. In the tropics the westerly bias in HadGEMminus is shifted southwards and intensifies and in southern mid-latitudes a large westerly bias of +6 ms^{-1} appears in the PS11 version which, while present, was much smaller in the previous model version. Figure 8 shows the zonal mean of zonal wind speed biases for the DJF season. At most levels and latitudes there is a clear reduction of the zonal wind biases for HadGEMminus apart from the strengthening of the westerly bias (to about 6 ms^{-1}) at around 80N and 400-500 hPa. This improvement reflects the reduced

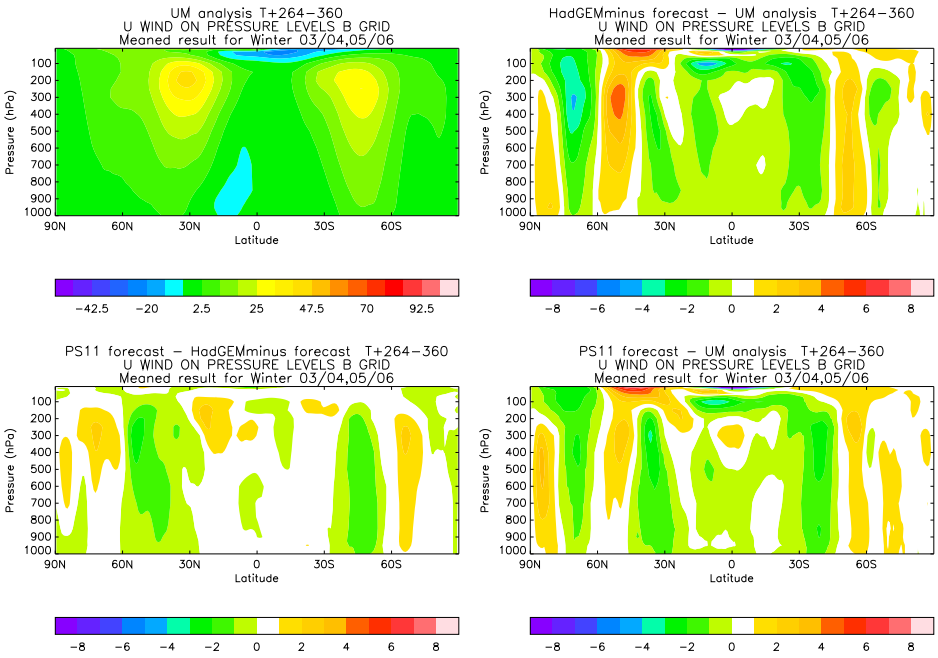

Figure 8: Zonal mean of u wind bias averaged over days 11-15 of all Winter cases

temperature biases seen earlier, while the strong westerly anomaly at 80N is consistent with the increased warm bias in northern mid-latitudes from the surface to around 600 hPa. In summer (not shown) the stronger cold bias in the southern polar regions at around 400 hPa causes a westerly bias at around 300 hPa.

Meridional winds

The zonal mean errors of the meridional winds averaged over days 11-15 for the summer cases are shown in Figure 9. This shows that there are only minor changes to the errors in meridional wind from the introduction of the new configuration with similar patterns of errors in the meridional winds and the errors having very similar sizes in both. Both versions are consistent with the findings of Milton *et al* (2007) of a complex error pattern with a strengthening of the circulation aloft but also a confining of the upper level outflow to a narrower height range. Although the T+24 meridional winds were shown to be improved by the modifications to convection introduced at PS10 (Willet, 2006) the changes do not seem to have a significant impact on the errors at longer timescales.

3.4 Systematic errors in the extra tropical circulation 500 hPa geopotential height.

Figure 10 shows geopotential height of the 500hPa surface in the winter for both hemisphere (i.e. DJF in NH and JJA in SH). The two winters have been chosen as the largest differences seen for this season. There are significant changes in circulation in both hemispheres with changes by -12 to +12 dam. The RMS error shown on plots of horizontal levels in this document is a measure of the magnitude of the systematic error: the 'spatial' RMS of the mean bias: $Spatial\,RMS = \sqrt{<\epsilon^2>}$; where$\epsilon = mean\,bias$, <>= area weighted spatial mean. There is a slightly lower 'spatial' RMS error in the southern hemisphere (3.42 not 3.79) but the RMS error increases in the northern hemisphere (from 2.92 to 3.24).

3.5 Global energy cycle

This section investigates the modelled global energy cycle as formulated by Lorenz (1967). The partitioning of energy into available and kinetic energy stores and then again into zonal and eddy components is calculated from

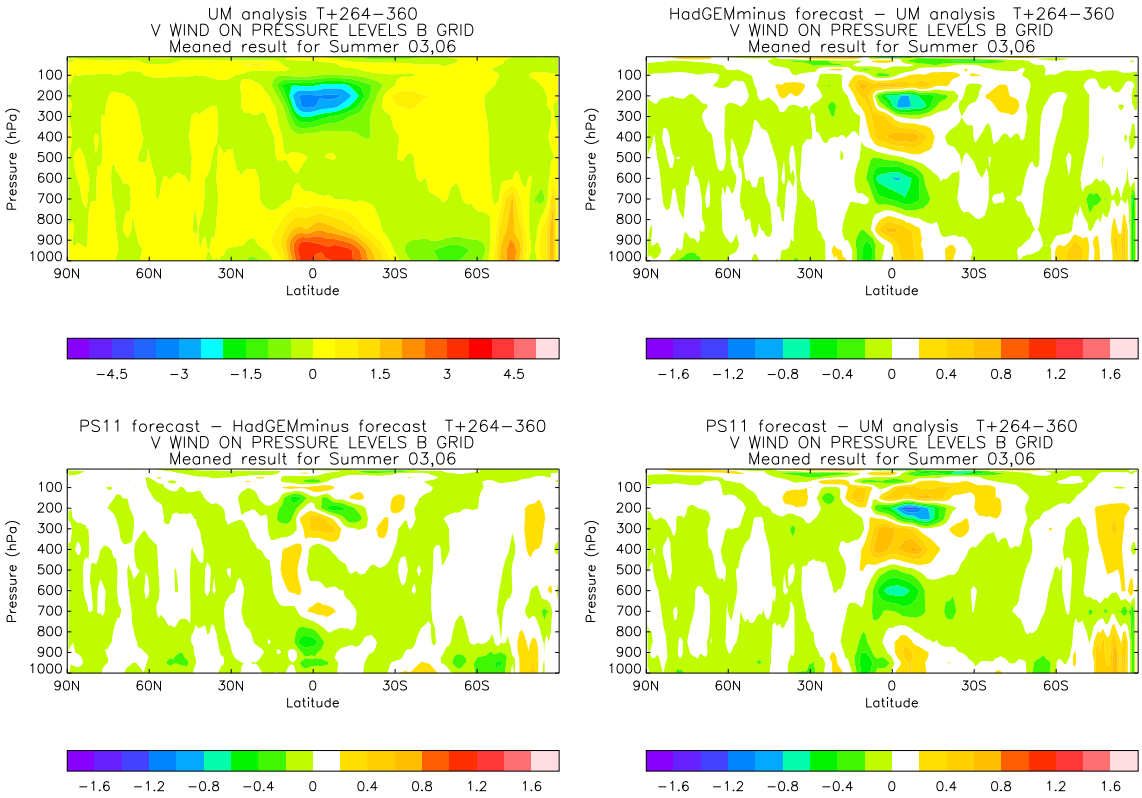

Figure 9: JJA errors in meridional winds from days 11-15

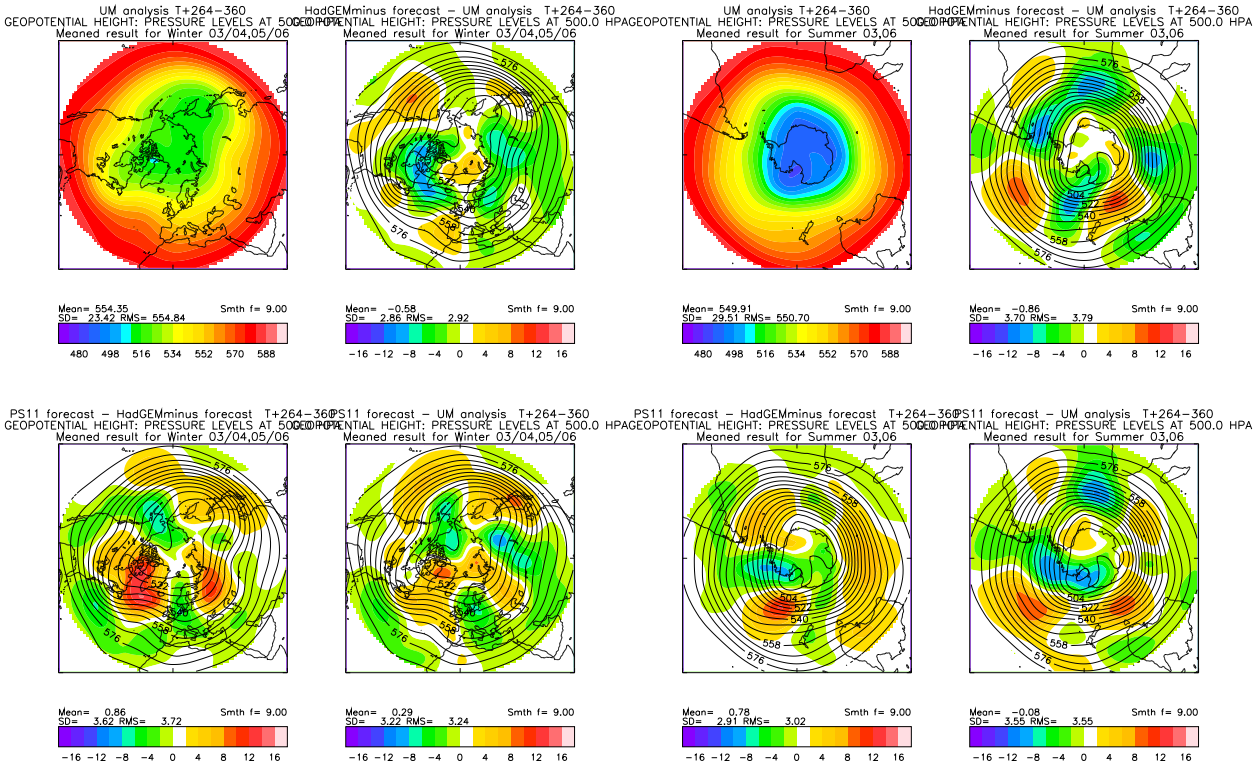

Figure 10: Geopotential heights at 500 hPa averaged over days 11-15 for both forecasts in hemispheric winter. Over-plotted contour lines on difference plots are geopotential heights from the analyses.

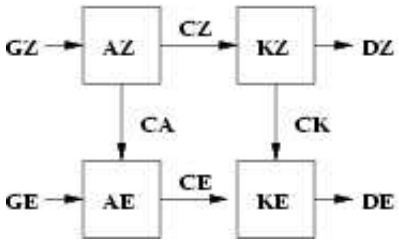

Figure 11: Lorenz Energy Cycle

both analyses and forecasts (fig. 11). This global energy cycle provides a useful check on the modelled variability compared to analyses in important terms such as the eddy kinetic energy. This discussion concentrates on the storage terms (AE, AZ, KE and KZ in figure 11) and the energy conversion terms are not discussed in any detail.

Figure 12 shows the energy storage terms in analyses and forecasts as daily averages of the JJA cases. Compared to Milton *et al* (2007) the energy budget terms in the analyses show smaller variations due to the increase in the number of cases from 4 or 5 to 10 for each season. However there is still some variation so it is important to compare the modelled energy budget terms to the analyses. For the summer case (JJA) available zonal potential energy (AZ) and zonal kinetic energy (KZ) increase in both forecasts. The analyses have a more constant value of AZ than the forecasts. The two forecasts give a very similar rate of increase of AZ until about day 7.5 then the HadGEMminus configuration peaks earlier and starts to fall to values closer to analyses at shorter forecast ranges than PS11. The increase in AZ was seen previously and was explained as a result of the cooling in the mid latitudes and warming in the tropics causing increased zonal winds through thermal wind balance considerations. In contrast to Milton *et al* (2007) both forecasts agree well with the analysis for the zonal kinetic energy term (KZ) with a small increase over the 15 day period so it seems likely that the increase in the number of cases has led to a better forecast - analysis comparison. The forecast eddy available potential energy (AE) shows a similar increase relative to analysis to AZ but the differences between the two model formulations are very small at all forecast ranges as for KZ. Eddy Kinetic energy (KE) shows a similar evolution to that seen in Milton *et al* (2007) with an initial fall off in KE in the forecasts relative to the analyses. This loss of eddy kinetic energy may be related to model representation with some eddies being too small to be represented in the model. After 5 days both forecasts increase steadily albeit at different rates. This increase at longer forecast ranges may be due to energy propagating via the baroclinic life cycle (AZ →AE→KE). HadGEMminus increases more rapidly and by about day 10 has more eddy kinetic energy than the analyses whereas the PS11 model formulation increases more slowly and by day 15 has the same values of KE as the analyses. The more rapid increase of KE in the HadGEMminus run may be due to the bigger temperature biases in these runs. In the DJF season (figure 13) the modelled energy cycles are in much better agreement with analyses than for the northern hemisphere summer consistent with the finding of Milton *et al* (2007). The two forecasts also show smaller differences from each other with the notable exception of zonal kinetic energy in which the summer forecasts agreed well with one another whereas in winter by day 15 the PS11 run is now about 20 kJm^{-2} smaller and in better agreement with the analyses.

3.6 Systematic errors in tropical circulation - divergent flow

Figures 14 and 15 show the bias in velocity potential at 850 hPa and 250 hPa averaged over days 11-15 for both seasons. For the northern hemisphere summer cases the biggest change is the decreased low level convergence and upper level divergence over the Indian Ocean in PS11 giving much better agreement with the analyses in this region although there is still too much convergence at low levels in this region. Low level divergence over South America is also decreased causing the model to change from having a low level divergent bias in this region to a convergent bias centred on Central America. The spatial RMS error falls at both levels. These differences are those which would be expected from the impact of adaptive detrainment. In DJF a similar improvement is seen in the Indian Ocean region. The changes over South and Central America are however much smaller and have less of an impact. Overall it can be seen that the tropical divergent circulation in PS11 is much improved compared to the earlier model configuration.

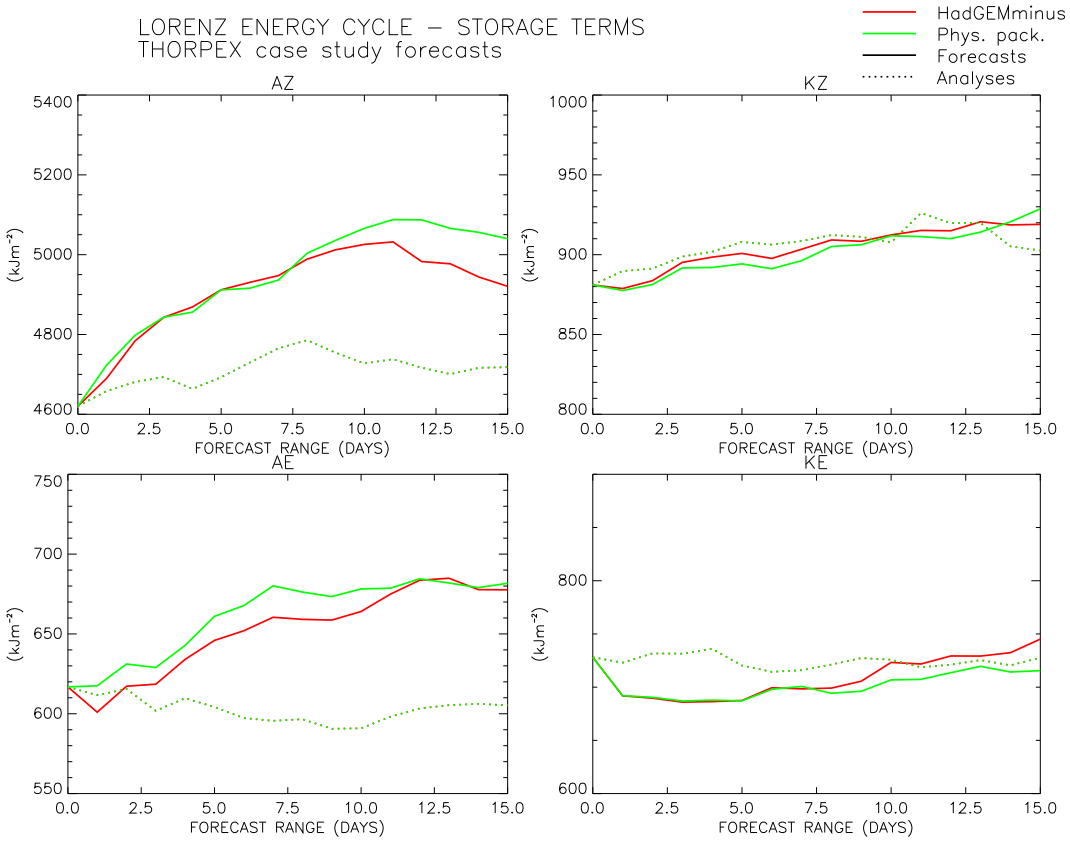

Figure 12: Lorenz energy cycle - energy storage components for JJA case studies

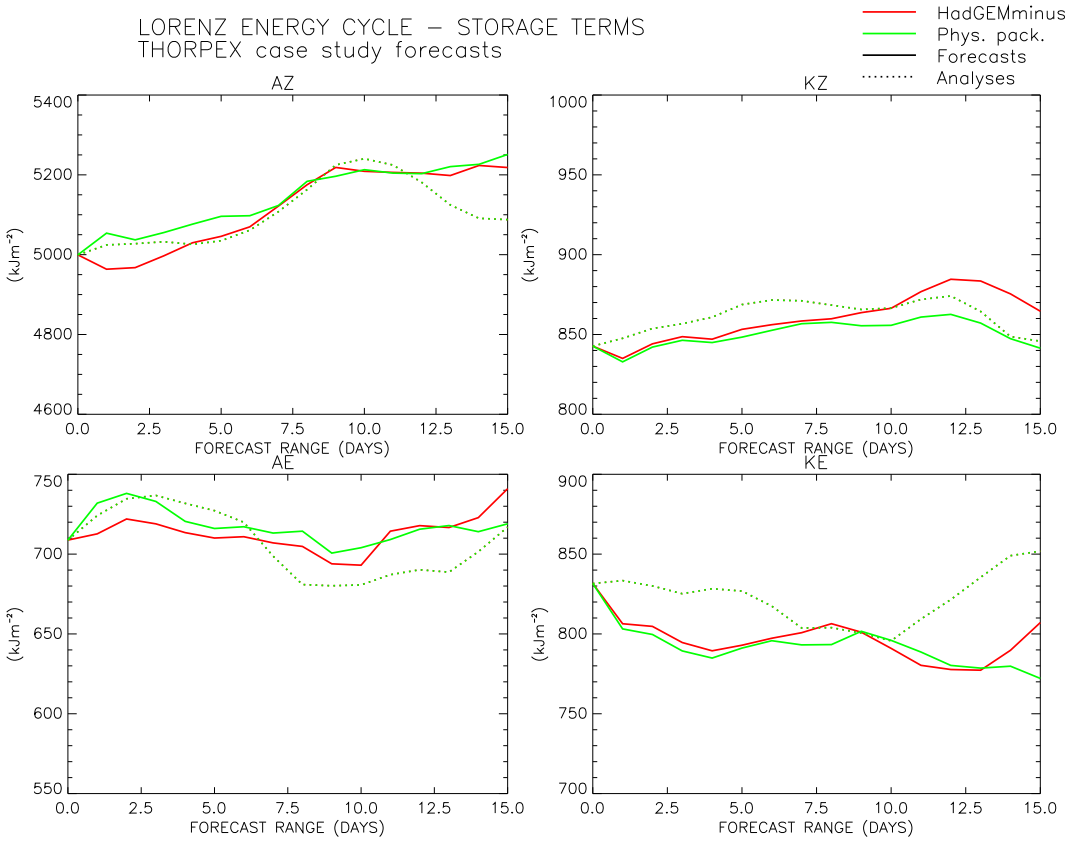

Figure 13: Lorenz energy cycle - energy storage components for DJF case studies

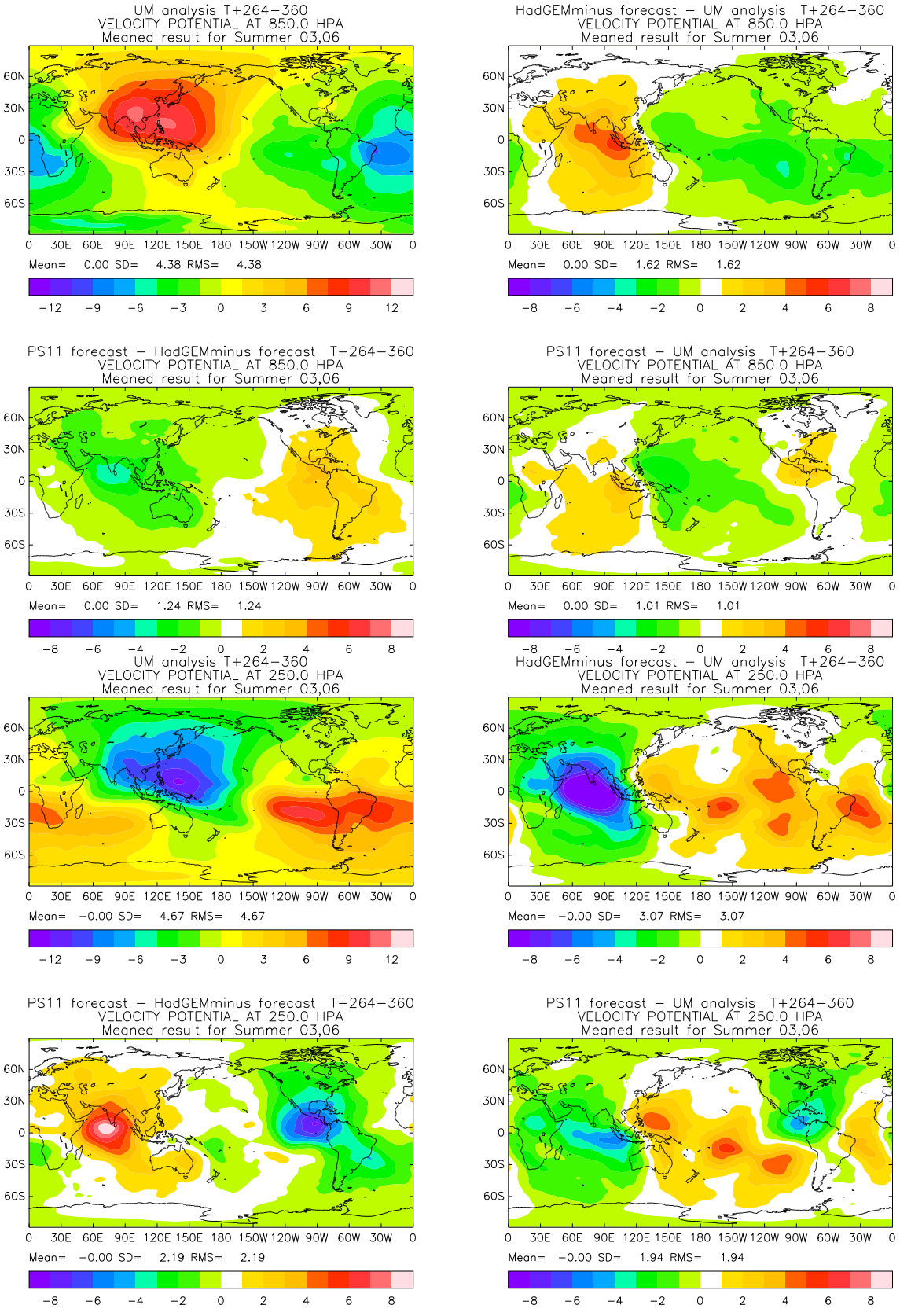

Figure 14: Velocity potential JJA at 250 and 850 hPa

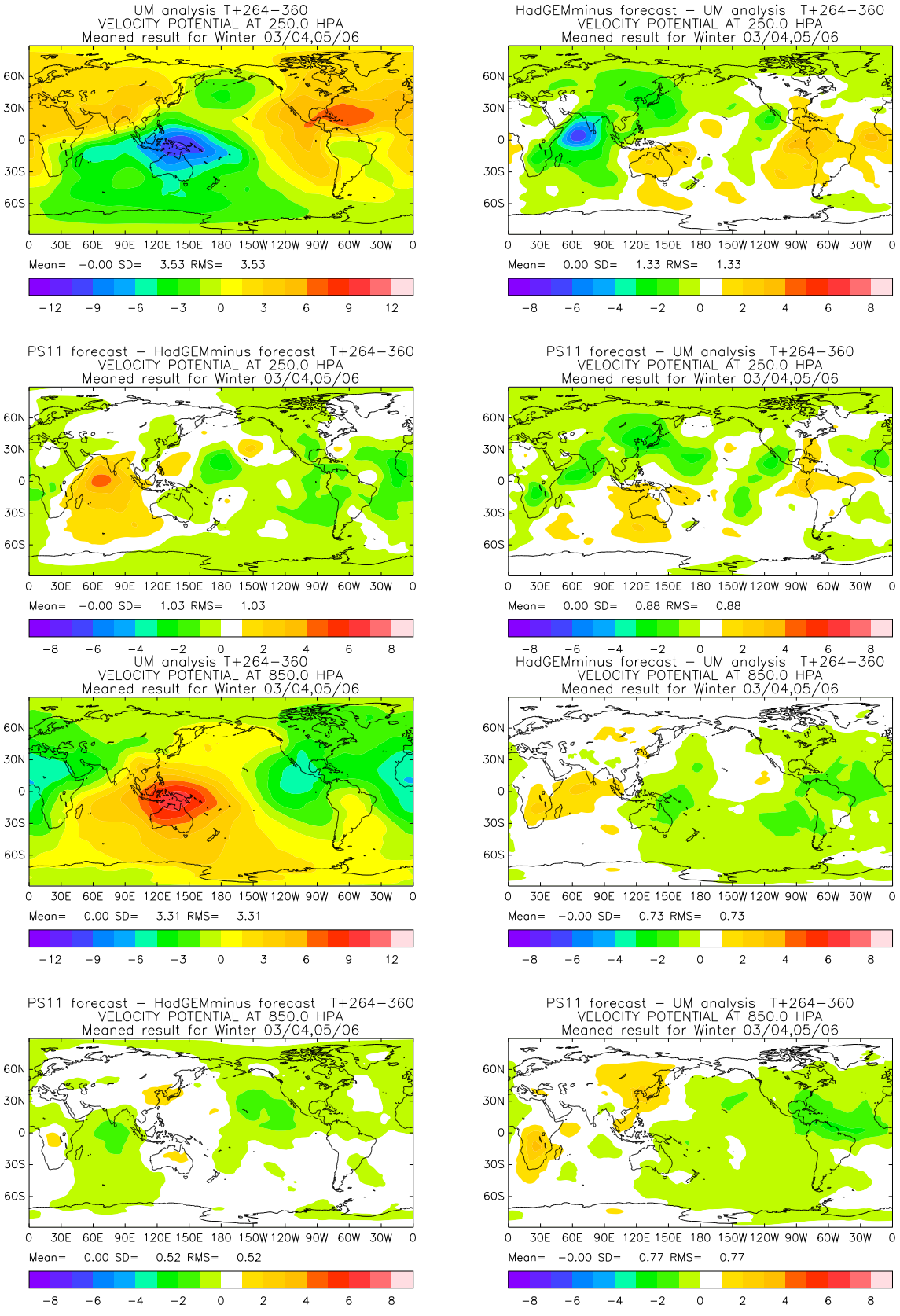

Figure 15: Velocity potential DJF at 250 and 850 hPa

3.7 Precipitation

To validate model precipitation use has been made of the GPCP[3] 1 degree daily (1DD) dataset (Huffman *et al* 2001). This provides daily data at 1x1 degree resolution on a global grid from January 1997 to close to the present (as of February 2007 the dataset included up to and including data for December 2006). The data set merges data from infrared and microwave satellite instruments with rain gauge data from over 6,000 stations. To evaluate the precipitation fields in the two model configurations 24 hour precipitation accumulations have been averaged from T+1.5 to T+14.5 and compared to the average GPCP daily precipitation over the same period (figure 16). This averaging period was chosen because the GPCP daily datasets start at 0UT. Note that unlike the data used in Milton *et al* (2007) GPCP 1DD data was available for all latitudes during this period.

For NH winter, there are many areas of improvement in the precipitation compared to GPCP. The changes are dominated in the tropical regions by changes in convective rain. In the Indian ocean there is a reduced positive bias (which appears consistent with the weaker ascent seen in velocity potential fields in section 3.6), there is smaller dry bias over North Australia giving a better representation of the Australian Monsoon which can also be linked to changes in the circulation seen above. However the precipitation is still underestimated over Australia and overestimated for Indonesia. In the mid Pacific there is a small reduction in the positive bias. Over many land regions the new physics has clearly had some negative impacts on the precipitation: it has made worse the positive bias over Africa and the north of South America. These biases are again similar to some were seen in the operational model. The physics modifications tested here, lead to a reduced cloudiness over land, causing an increase in downward shortwave radiation. This drove a more intense hydrological cycle. To improve the modelled interaction of cloud and radiation, the operational model has now been changed so that the convective cloud amount which interacts with the radiation scheme undergoes an exponential decay with a lifetime of 2 hours. This improved these precipitation biases. There is a new small positive bias in the Arctic connected to increased cloudiness and an increase in large scale snowfall (not shown) but the GPCP data is unlikely to be highly accurate at these high latitudes. Overall the spatial RMS and mean error are similar between the two model configurations.

For NH summer similar changes are seen to those in winter, with the largest differences in the tropics. The reduction in the positive bias in precipitation over the Indian Ocean is stronger than in the winter season, as would be expected from the velocity potential changes. The negative impacts on precipitation over land areas are also present in this season.

3.8 Cloud cover

Figure 17 shows the total cloud cover averaged over days 11 to 15 for both the winter and summer forecasts compared to seasonal mean total cloud amounts from International Satellite Cloud Climatology Project (ISCCP). It can be seen that there are big increases in Arctic cloudiness in DJF giving much better agreement with ISCCP. This helps explain the removal of winter cold bias near the surface in these regions and a similar argument can also be applied in the north of Russia where decreased cloud occurs in a region where PS11 has lower 1000hPa temperatures. There are also large increases in cloud over North Australia with a decrease over Indonesia, giving improved agreement with ISCCP, which is consistent with the improved modelling of the monsoon in this region as seen in the precipitation. In northern hemisphere summer the decreased cloud cover over Western Europe can be seen to be linked with the reduction in the cold bias in this region but these lower values are in worse agreement with ISCCP. The decreased cloud in parts of N. America and Asia may be contributing to the warm bias there. There is also less cloud cover over much of the tropical oceans which is consistent with the finding of Willet (2006) on the impacts of the spring physics package. In both seasons the positive cloud bias over the Indian Ocean is reduced which is also consistent with the changes in the precipitation and velocity potential already discussed.

3.9 Grosswetterlage

The Grosswetterlage (GWL) catalogue is a fixed set of typical weather patterns used to classify synoptic weather regimes in Europe. The method used here is based on the daily catalogue developed by Hess and Brezowski and maintained by the German Weather Service (DWD). James (2006) has adapted the original methodology to develop an objective method for assigning each day to a particular weather type. The 29 GWL regimes are large-scale

[3]Global Precipitation Climatology Project

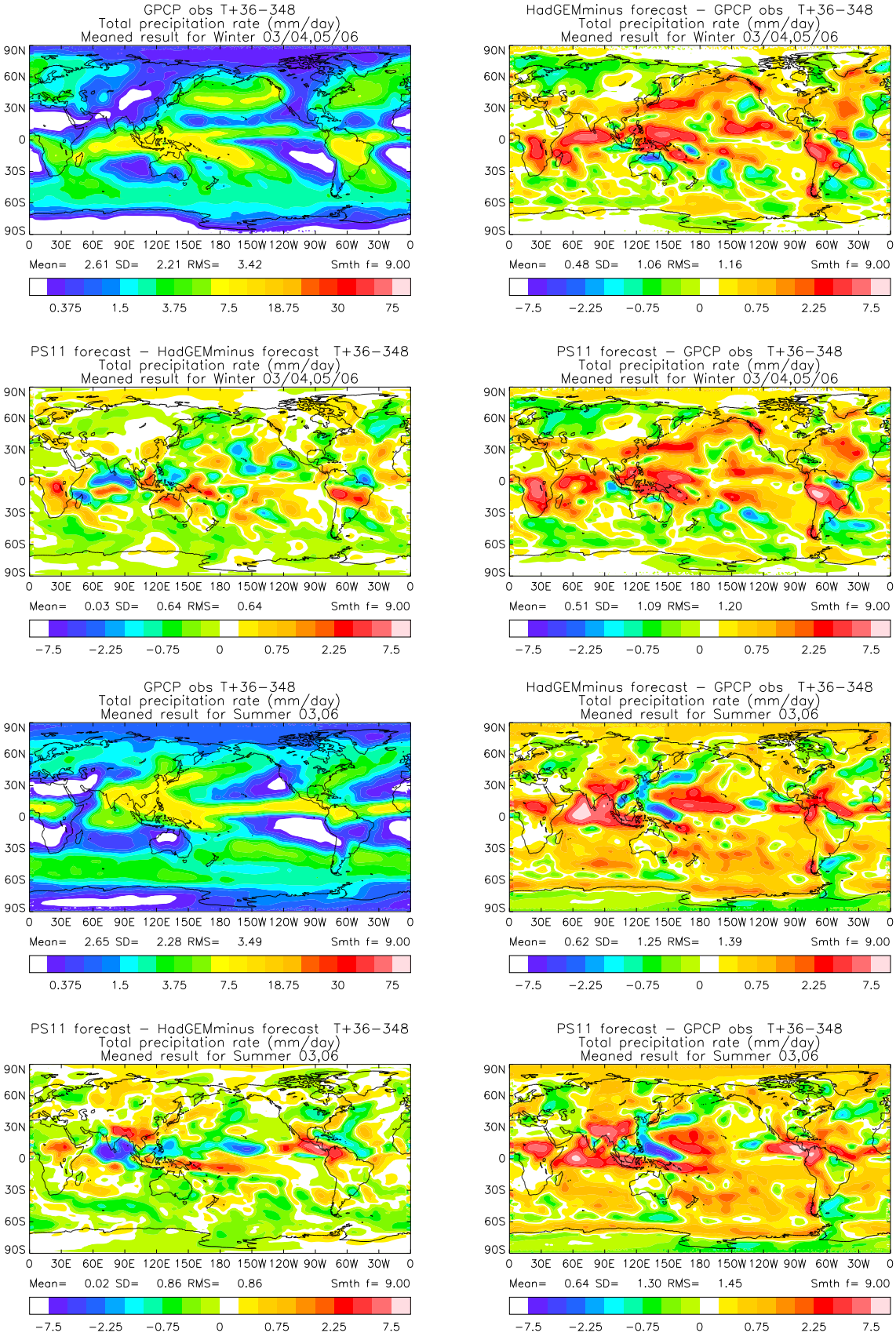

Figure 16: Model precipitation for Northern Hemisphere winter (top) and Summer (bottom) versus GPCP daily precipitation data.

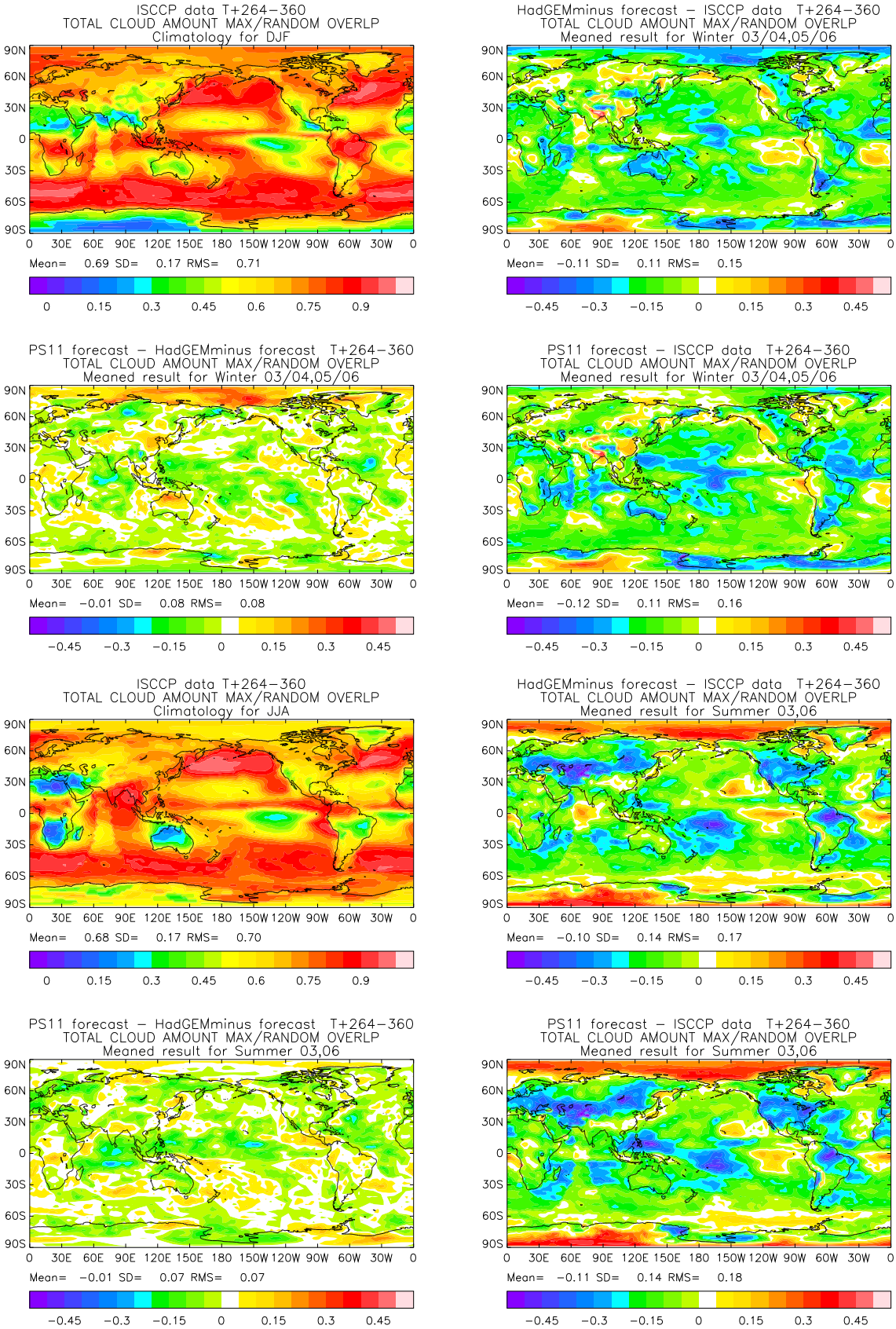

Figure 17: Total cloud cover averaged over days 11-15 for DJF (top) and JJA (bottom) versus ISCCP data.

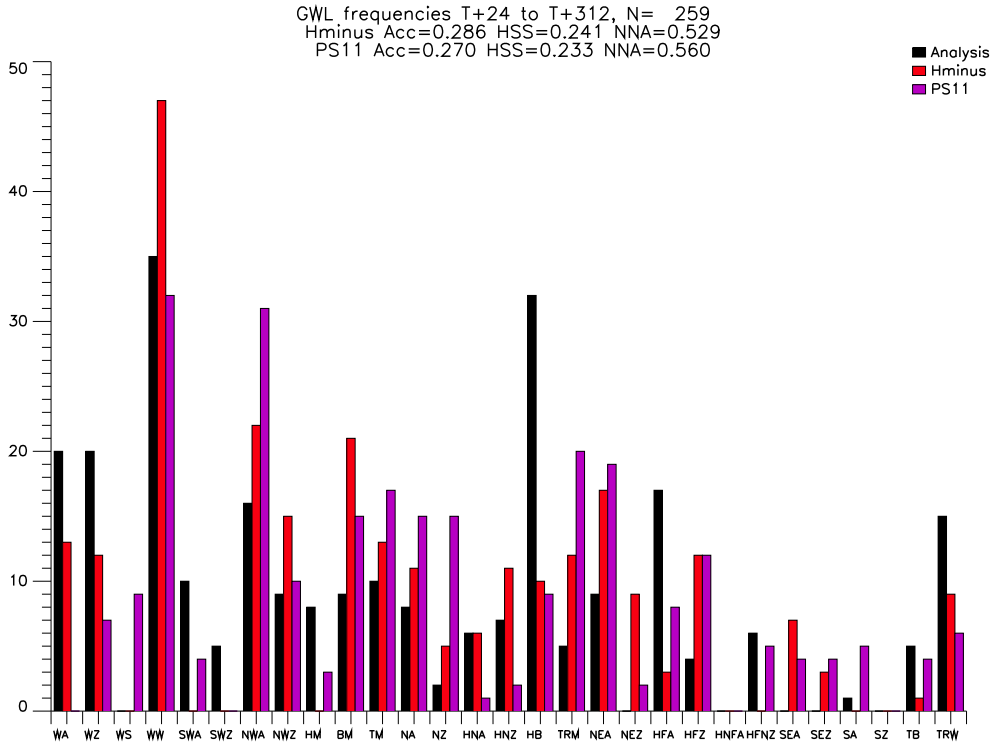

Figure 18: Distributions of GWL types in the 2 model configurations and skill scores. Data from all forecast ranges was used. Black= ERA-40 analysis, red =HadGEMminus, purple=PS11.

circulation patterns involving the whole of Europe and the North-East Atlantic, with their primary focus on central Europe and aim to relate directly to weather as actually experienced in Europe.

The GWL type for days 1-13 of each forecast have been calculated (the method requires a given type to persist for a minimum of 3 days so the GWL categories for the last 2 days of the forecast could not be calculated) and compared to a catalogue of GWL types calculated from ERA-40 reanalysis. Figure 18 compares the marginal distributions of the 2 forecasts to ERA-40. It also gives for the two forecasts the forecasts accuracy (Acc), Heidke Skill Score (HSS) and Nearest Neighbour Accuracy, (NNA). NNA is defined as the proportion of forecasts for which the model predicts the correct GWL type or one of five weather regimes which are the most similar to it. The results here for PS11 appear to be worse with both a lower Heidke Skill Score (HSS) and an increased bias away from Westerly air flow regimes and into more Northerly types. However, due to the large number of categories and the limited number of forecasts it is not clear if this is a statistically significant result. The PS11 forecasts also have a slightly higher Nearest Neighbour Accuracy but again this may not be statistically significant. GWL types are being forecast routinely by the MOGREPS-15 system and these statistics are being compiled for those results. When a larger database of results from this system has been compiled more significant results should be possible.

3.10 Blocking

Blocking is a key process in creating extremes of weather at mid-latitudes and provides a useful test of the model's ability to represent the atmosphere. To evaluate the models' ability to forecast blocking we use a blocking index based on that of Tibaldi and Molteni (1990) as modified by an additional constraint relating to the gradient in geopotential height suggested by D'Andrea et. al. (1998). At any given longitude, blocking is diagnosed based on geopotential height gradients indicating the presence of geostrophic easterlies in the region of the mid latitude storm track. In addition the definition was modified following Pelly and Hoskins (2003) to use a variable latitude of the storm track. Blocking indices were calculated for all 20 cases (winter and summer) to increase the amount of data available.

Figure 19 shows the results of the calculations of blocking frequencies for the days 1-15, 1-5, 6-10 and 11-15 are shown in separate panels. The frequency of blocking in the analysis for each time period is shown in black. The

frequency of correctly predicted blocking (hits) are shown in red, the frequency of events where blocking was predicted but not observed (false alarms) in blue and the sum of false alarms plus hits is shown in green. For the whole 15 day period the maximum number of blocking events is 300 (15 days x 20 cases) and for each 5 day period the maximum is 100. The solid coloured lines are the results from PS11 and the dashed lines are from HadGEMminus. A perfect model would have no false alarms and the frequency of hits (red lines) would equal the frequency of blocking in the analysis (black line). If the average frequency of blocking in the model were correct but the predictability was not perfect then the sum of false alarms and hits (green lines) would equal the analysis (black lines).

For the Northern hemisphere, over the whole fifteen days (top left panel) the total frequency of blocking in both model versions agrees fairly well with the analyses, however there is too little blocking between 60-100 degrees and too much from 300-360 degrees indicating a westward shift of blocking which has been seen in operational forecasts at days 1-5. However a very large proportion of the blocking events are false alarms indicating a lack of predictability of blocking for these forecast ranges. At days 1 to 5 the frequency of false alarms is low in both model versions and, although the frequency of blocking in the Pacific sector is somewhat under-predicted, in general the results for both model versions are good and there is no evidence of the westward shift of blocking in the Atlantic sector. However at 6-10 days, although the overall frequency of blocking agrees reasonably well with the observed frequency, the frequency of false alarms is equal or greater than the frequency of hits, indicating very little predictive skill at this forecast range. There is still little difference between the two model versions and the westward shift of the region of Atlantic blocking can now seen. By days 11-15 the frequency of false alarms is often greater than the frequency of hits, with a large number of false alarms appearing between 330-360 degrees, indicating a westward shift of the Atlantic blocking and some indications of a Westward shift of the Pacific blocking region with the peak in false alarms to the east of the peak in observed blocking events. In the Southern Hemisphere the model performance for days 1-5 is similar to the Northern Hemisphere with few false alarms, a good agreement on the overall frequency and regions of blocking although with a sightly smaller frequency in the forecasts than in the analyses. At days 6-10 however there still seems to some skill in predicting blocking in the main region although the the overall frequency has fallen further but still with a low proportion of false alarms in this sector. However again by days 11-15 the false alarms are more frequent than the hits for many regions. Finally there is no clear improvement in skill for PS11 in the Southern Hemisphere - although at days 6-10 PS11 has more hits and a higher overall frequency of blocking at days 11-15 the situation is reversed with HadGEMminus having more blocking. Any differences between the two model versions for these 20 cases are unlikely to be statistically significant.

3.11 Madden Julian Oscillation

The Madden-Julian Oscillation (MJO, for more details see review articles in Lau and Waliser, 2005) is a key phenomenon for prediction on sub-seasonal timescales but has so far proved a difficult challenge for GCMs (Slingo *et al* in: Lau and Waliser, 2005). The MJO can be characterised as eastward propagating regions of deep convection and suppressed convection anomalies in tropics across a region from the Indian Ocean to the western Pacific with a period of around 40 days. It is associated with global-scale wind anomalies and was first detected by spectral analysis of zonal winds in the mid and upper troposphere. It has also been shown to have an effect on the extra-tropical circulation especially in the Pacific and Americas. To examine how well this phenomenon is captured in the two model configurations used here the MJO index of Wheeler and Hendon (2004) and the daily mean precipitation between 20N and 20S are examined.

Figure 20 shows the Wheeler and Hendon MJO index for the period of the 5 forecasts in DJF 2003-4. The Wheeler and Hendon index projects the daily anomalies of 3 fields (U850, U200 and Outgoing Long-wave Radiation) onto the 2 leading EOFs of the anomalies in the tropics (30N to 30S). The Real-time Multivariate MJO series 1 (RMM1) is the time series of the first principle component, and RMM2 is the second PC time-series . EOF1 corresponds to a situation with enhanced convection over the maritime continent and low level westerly wind anomalies over the Indian Ocean and Maritime continent. The second EOF lags EOF1 by 1/4 cycle and has enhanced convection over the Pacific. If observed values of RMM1 and RMM2 during an active MJO are plotted as in Figure 20 the observations are seen to follow an anti-clockwise path through the RMM1-RMM2 phase space. The strength of the MJO is given by the distance from the origin ($\sqrt{RMM1^2 + RMM2^2}$) and the phase of the MJO is given by the sector of the plot. The plots below are labeled with the regions representing the area where convection is heightened. The path taken by the analysis during this period shows how an MJO develops during the period of the first forecast with an initial growth in amplitude for the first ten days and then eastward propagation of the deep convection. This pattern is replicated in both forecasts although after the first five days both forecasts

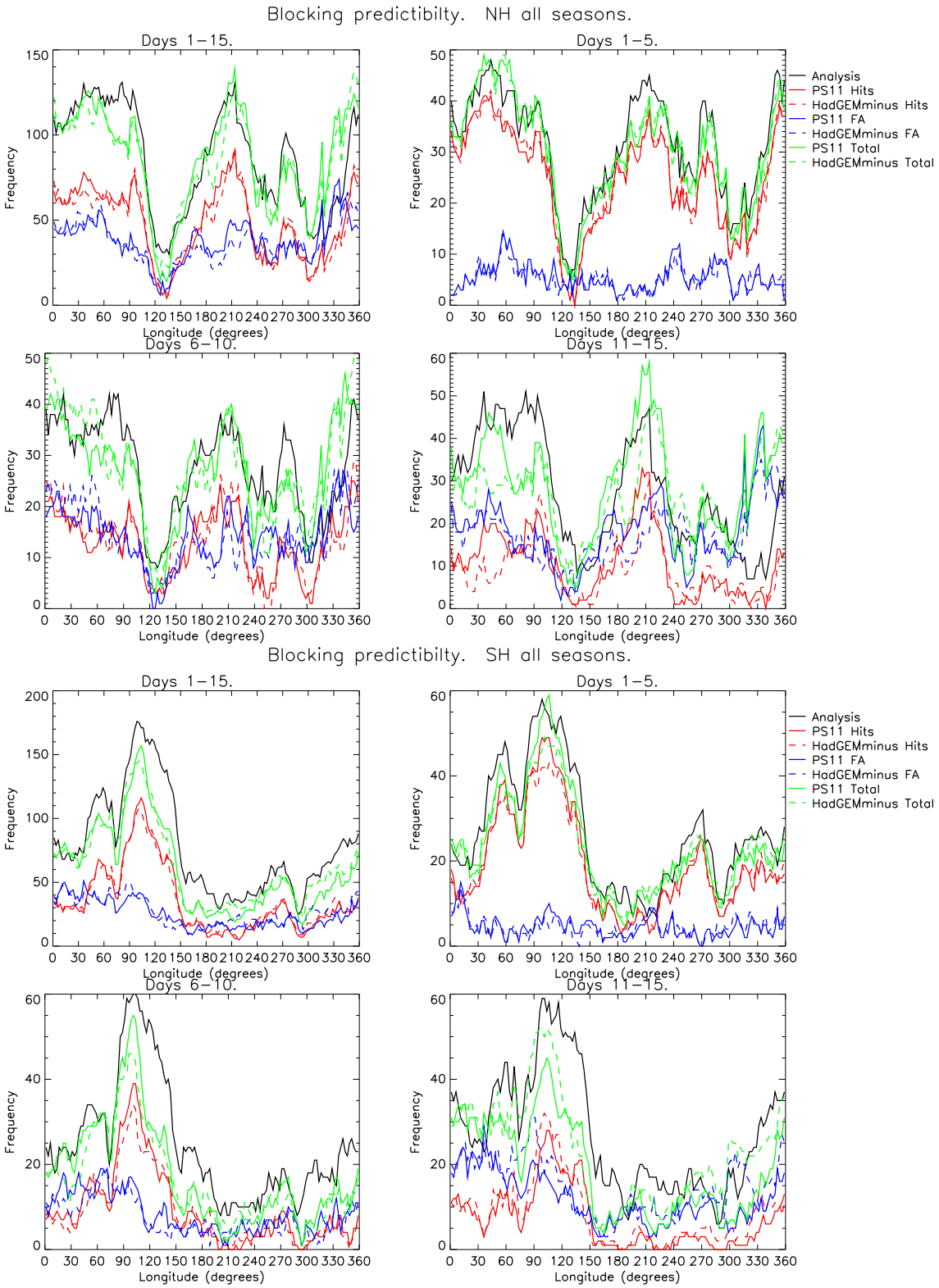

Figure 19: Blocking predictability in Northern Hemisphere (top) and Southern Hemisphere (bottom). Black-analysis, dashed lines HadGEM minus, solid lines PS11. Red lines= hit frequency, Blue line= False Alarms, Green lines= Hits+False Alarms

	HadGEMminus	PS11
RMM1	0.76	0.69
RMM2	0.92	0.72
Amplitude	0.68	0.55
Phase angle (degrees)	51	41

Table 2: RMS error over all 20 cases for both model versions of RMM1, RMM2, amplitude and phase angle.

have a smaller value of the RMM1 index putting them in sector 2 not 3 as for the analyses. The next 3 forecast periods show a good example of the MJO with a steady counterclockwise motion of the indices in phase space and completing more than a full MJO cycle and then decaying to give near random motion near the centre in the final forecast period. Both forecasts reproduce reasonably well this behavior, in general occupying the correct sector of phase space until the end of the fourth forecast where they both get stuck around the Indian Ocean region and the region of deep convection fails to propagate into the Maritime continent as seen in the analysis. There is not much visible difference in the MJO index between the two forecasts, although the PS11 forecast generally appears closer to the analysis at the end of the 15 day forecasts. Table 2 shows the root mean square error of the two indices RMM1 and RMM2 and for the amplitude and phase angle for both model versions calculated over all twenty forecasts. Note that the RMS error for the angle is only calculated for those occasions when the amplitude of the analysis is greater than 1. The RMS errors for the new model version are lower for all 4 parameters.

Figure 21 is a Hovmoller plot of the daily precipitation in the tropics from GPCP data and the two model configurations covering the same period as the MJO indices discussed previously.

Consistent with the MJO indices shown above two MJO cycles can be seen in the GPCP data. One begins at around 90E in the Indian Ocean on the 6th December and the precipitation maximum propagates eastwards into the central Pacific (180 degrees) over a period of 20-30 days. The precipitation becomes weak over the Eastern Pacific but then re-intensifies over Central/South America on 12th Jan. The second cycle beings slightly further East at about 60E on 21st January and after becoming quiescent over the East Pacific again leads to active convection over South America on the 16th Feb.

The most noticeable difference between the model runs is the weakening of the spurious westward propagating disturbances seen in the HadGEMminus run which are not present in the GPCP. The first MJO cycle seems to also be better represented by the PS11 model forecasts especially in the first 15 days of the forecast when the MJO is over the Indian Ocean. However there is no evidence in either model of the MJO propagating across the West Pacific and becoming active again over South America as seen in the GPCP . This is in contrast to the MJO diagnostic above where the evolution of the MJO is reasonably well captured in 3 of the 4 forecasts which have a strong MJO. The MJO in general is more clearly seen in dynamical fields (see also Milton et al, 2007) which is indicative of problems with the coupling between dynamics and physics for equatorial modes.

4 Comparison of results to control of ECMWF EPS

In order to gain an understanding of the size of the biases shown here compared to those of other forecasting centres a comparison with results from the control of the ECMWF Ensemble Prediction System (EPS) has been carried out. Although the full ensemble was only run for 10 days during the period of the forecasts studied here, for research purposes the ensemble control was run out to 21 days and archived throughout the period of interest. The horizontal resolution until 31 January 2006 was T255 (~80km) and thereafter was T399 (~50km). The ECMWF operational analyses and EPS control forecasts have been meaned for the same periods as the model runs shown above. They have also been re-gridded to N144 resolution and then compared to the model run with new physics as described above (PS11). A more detailed discussion of systematic errors in the ECMWF forecasting system can be found in Jung and Tompkins (2003) and Jung (2005). These studies were carried out with many more forecasts than have been used here but were also performed with older versions of the model (the most recent model version used was that operational in 2003). As a result the systematic errors are not directly comparable between those studies and the results presented here. However, the magnitude of the errors has been found to be similar and for some fields the patterns of the errors are similar.

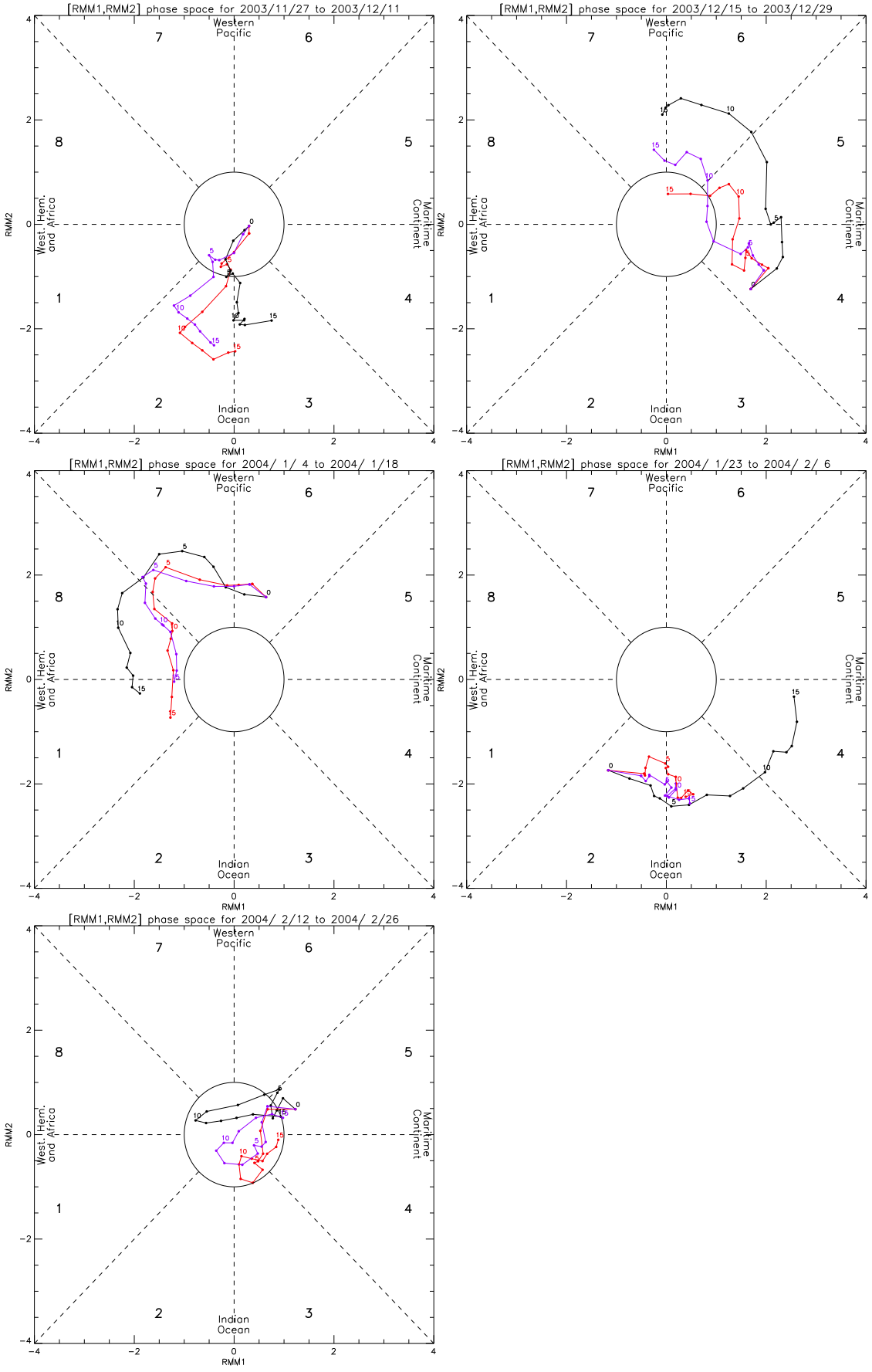

Figure 20: Plot showing analysed (black) and forecast (red - HadGEMminus, purple - PS11) MJO indices for period from 27/11/2003 to 26/2/2004.

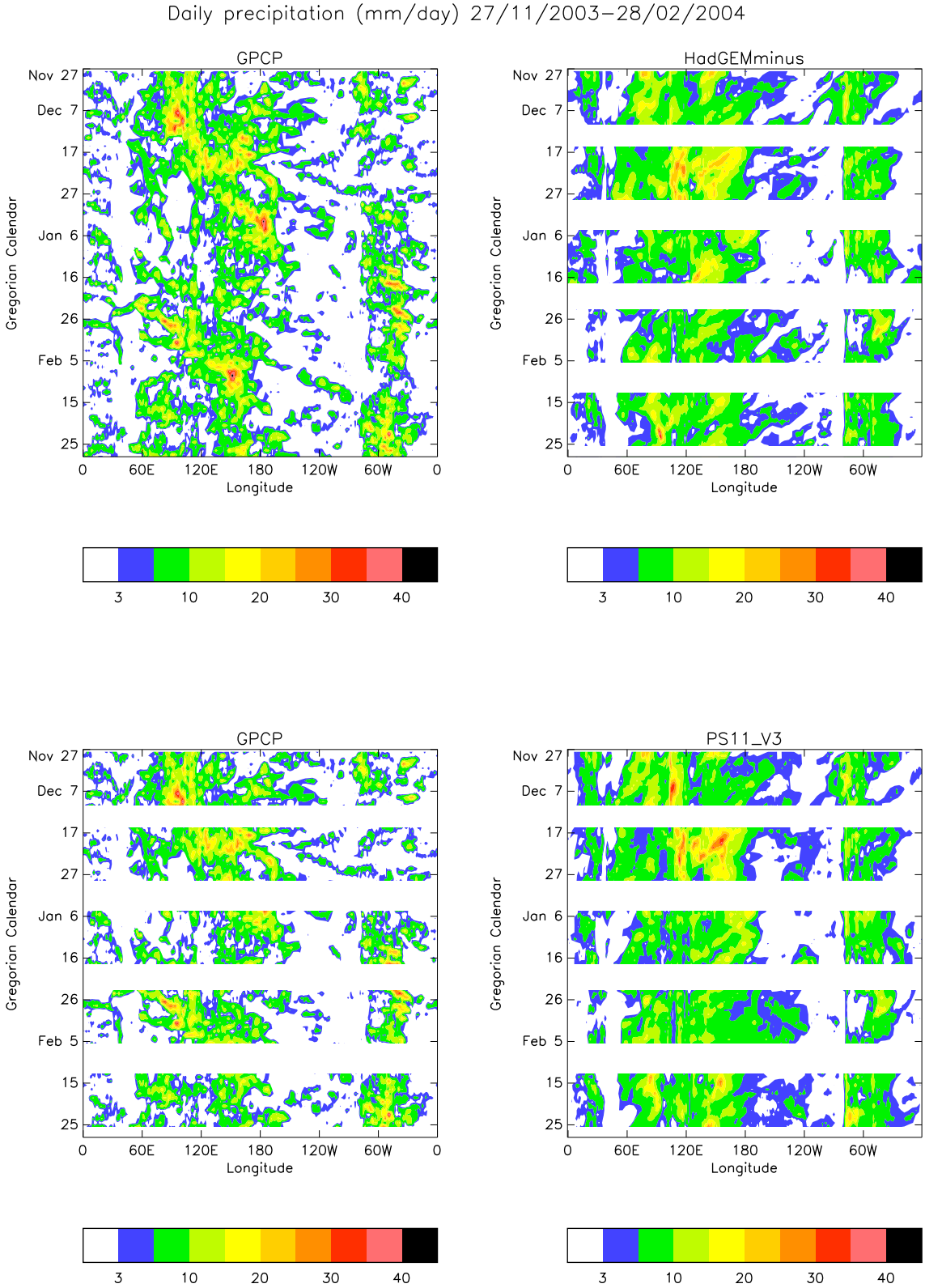

Figure 21: Hovmoller plots for 27/11/2003 to 28/2/2004 of daily precipitation averaged from 10N to 10S from GPCP and both model configurations.

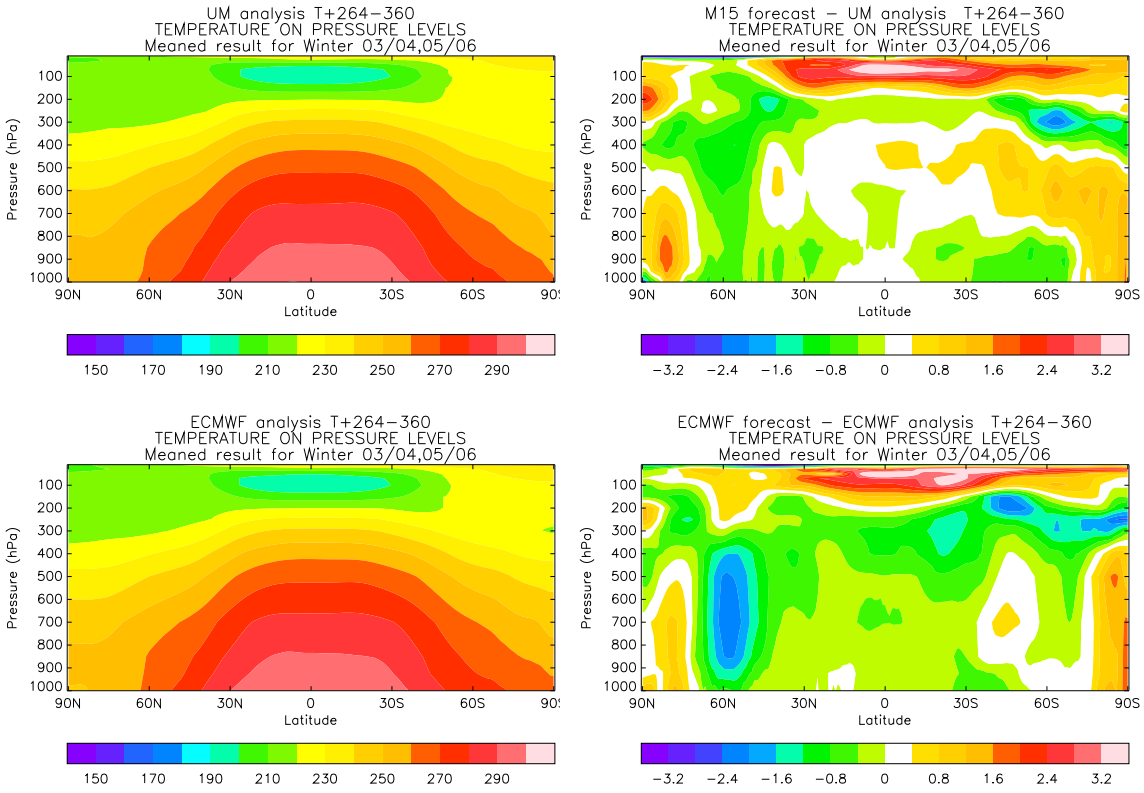

Figure 22: Comparison of zonal mean Temperature from PS11(top) and ECMWF ensemble control (bottom)

4.1 Results

For certain regions and fields the Unified model performs similarly or slightly worse, for others it is significantly worse and for some fields the Unified model may be slightly better. Some examples of each of these will be presented here. Figure 22 compares Temperature biases in the 10-15 forecasts for the DJF season. The left hand size plots are the UM analysis (top) and the ECMWF analysis (bottom). On the right are the differences of the forecasts from the respective analysis. Both models have a dipole type bias at 60N to 80N with a similar difference between the maximum and minimum but the UM has a higher maximum value and a higher minimum. Elsewhere the size of the errors are similar.

Figure 23 shows the same plot for zonal winds. Again the errors are of a similar magnitude although in several places the UM error are slightly larger. What is most interesting however, is that in most of the troposphere the patterns of the errors are remarkably similar, suggesting common issues with the two models. This agreement in the pattern biases is also seen for several other fields.

The JJA season 1000hPa temperatures are shown in Figure 24. The warm bias which is a known problem in the UM is present in both models, however it is much more severe over North America in the UM. The bias over Asia is still larger in the UM but the difference between the two models is less pronounced.

The meridional winds in DJF as shown in Figure 25 appear to be somewhat better in the UM, however the meridional winds may not be well constrained in the analyses and so these differences may not be significant.

Finally, Figure 26 shows the zonal mean of relative humidity for JJA. The main source of differences in the analyses is the different definitions of relative humidity used. In the ECMWF results RH is defined with respect to water above above 273.16 K, to ice below 250.16 K, and with respect to a mixture of both in between while the Unified model uses RH with respect to water everywhere. The forecasts also have very different error fields with a moist bias at 100hPa in the UM, and a dry bias below, while in the ECMWF model there is a dry bias at 100hPa and a moist bias below.

Overall these results are very encouraging, giving confidence in the performance of the MOGREPS system relative to ECMWF at least as far as issues related to biases are concerned. Although there are clear areas where improvements

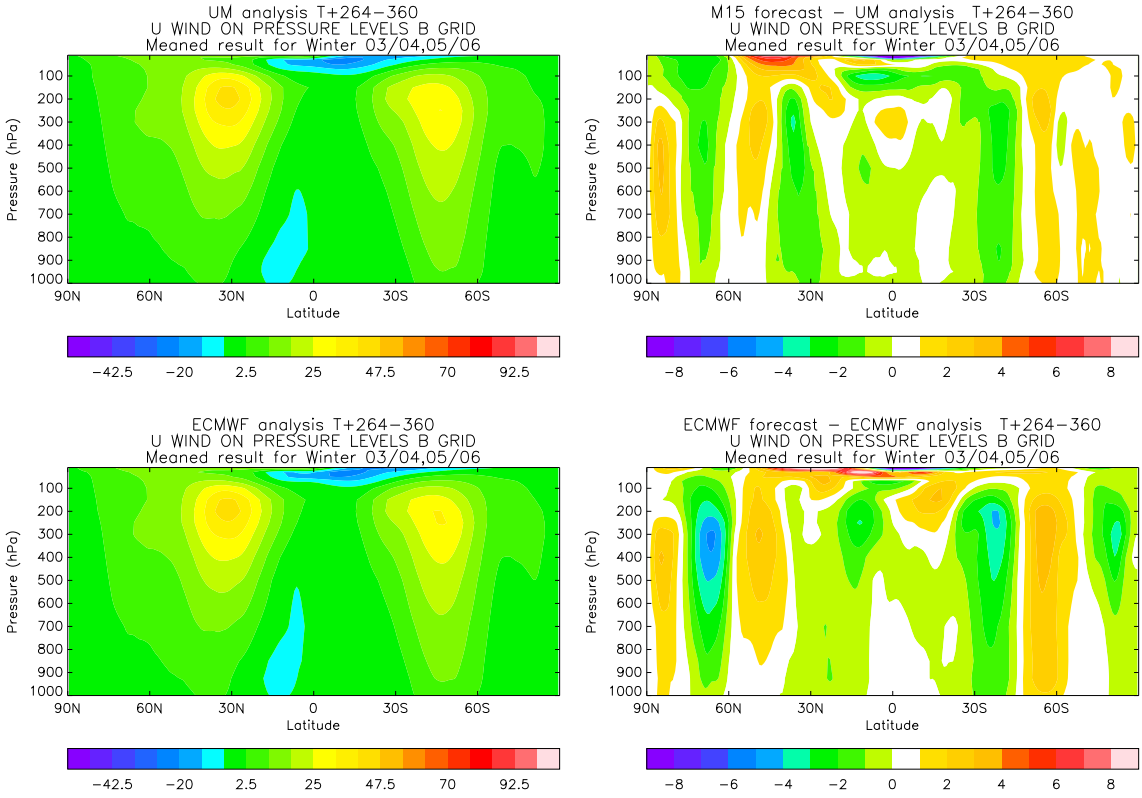

Figure 23: Comparison of zonal mean u winds from PS11(top) and ECMWF ensemble control (bottom)

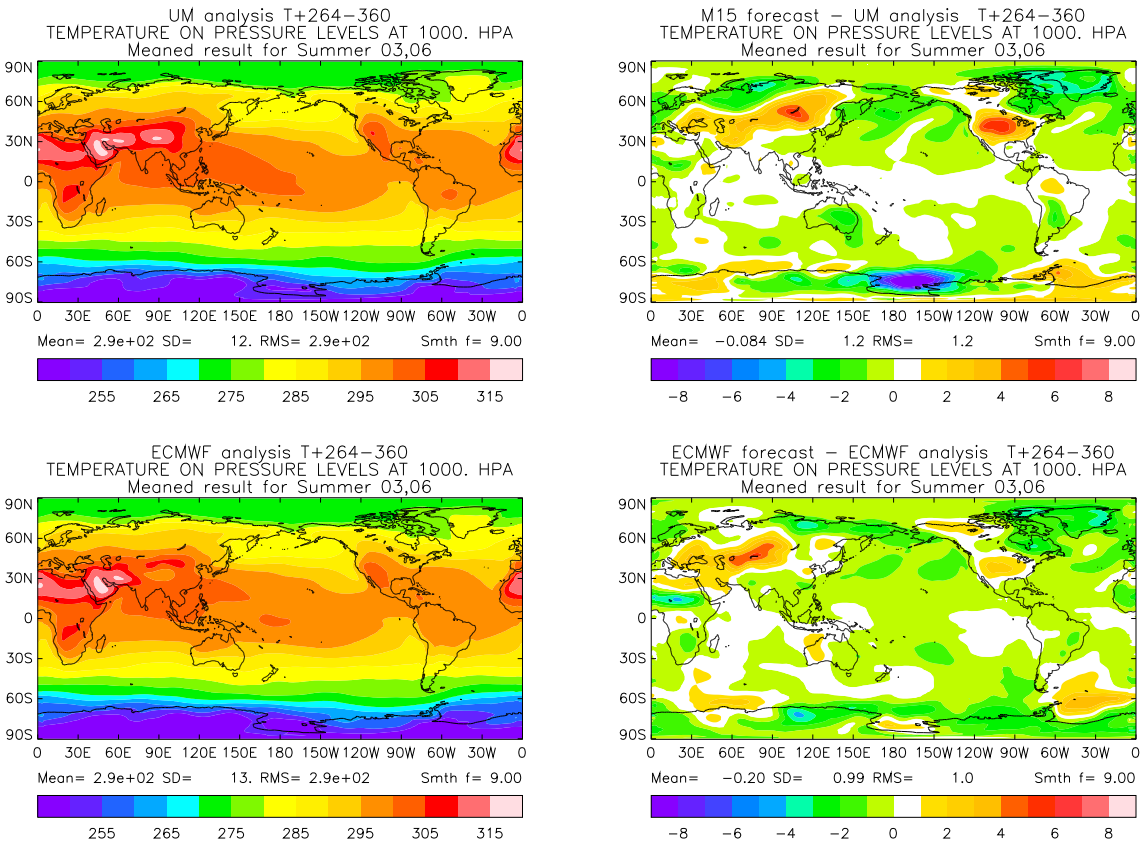

Figure 24: Comparison of 1000hPa Temperature from PS11(top) and ECMWF ensemble control (bottom)

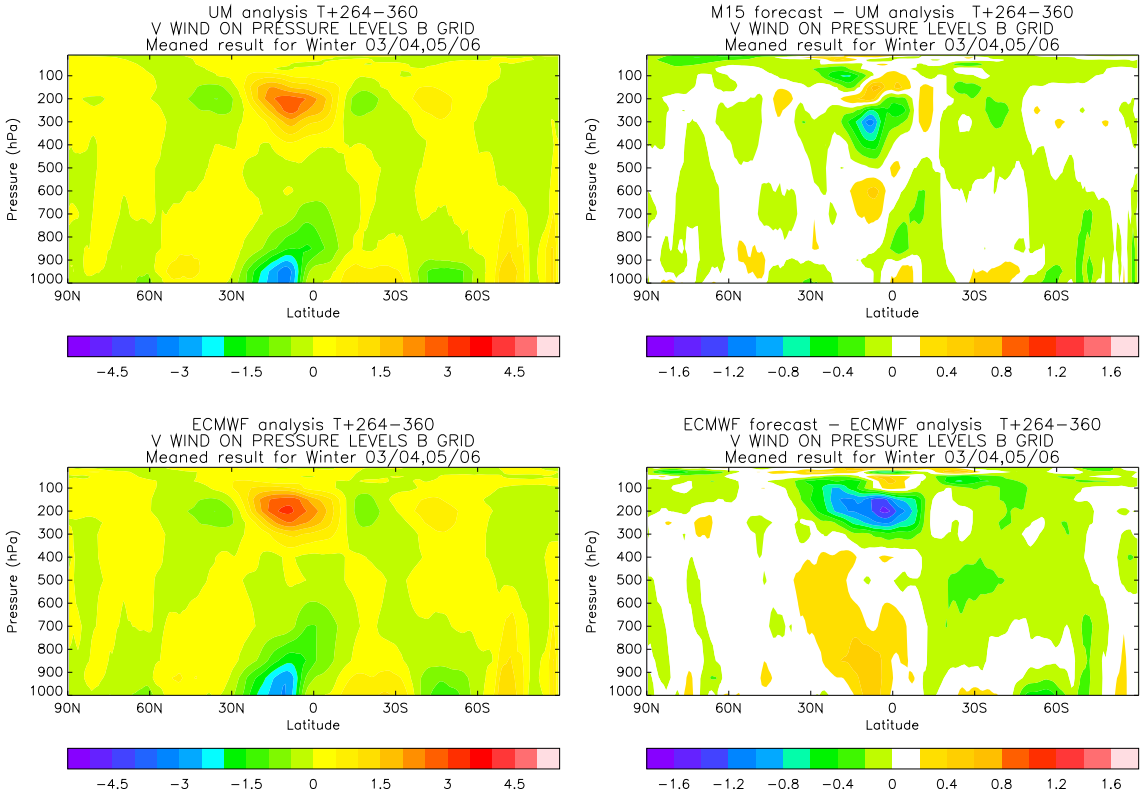

Figure 25: Comparison of zonal mean v wind from PS11(top) and ECMWF ensemble control (bottom)

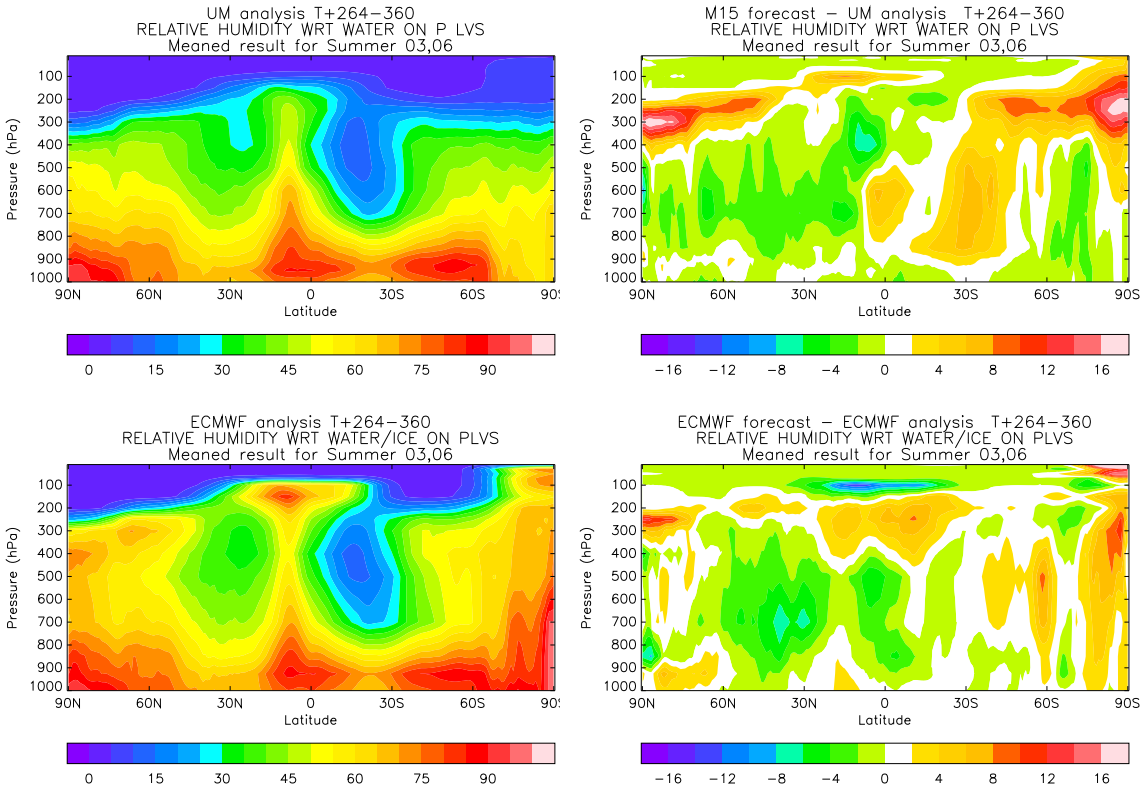

Figure 26: Comparison of zonal mean RH from PS11(top) and ECMWF ensemble control (bottom). Note that the Relative Humidity in the ECMWF results is defined with respect to water above above 273.16 K, to ice below 250.16 K, and with respect to a mixture of both in between.

can be made, the results suggest that the version of the system based on PS11 can provide a useful basis for 15 day forecasting.

5 Summary and Future Work

We would recommend the introduction of the the new physics package into the Met Office THORPEX ensemble prediction system. These changes have been shown to give large improvements in the model in the winter northern hemisphere surface cold bias and to the tropical divergent circulation. Other improvements are seen in monsoon precipitation over North Australia and precipitation in the Indian ocean and the representation of the Madden-Julian Oscillation. This latter result is in contrast to the results of Inness (2007) and Marshall (2006) for a climate version of the Unified Model - HadGEM1a (an updated version of HadGEM1) in which the strength of the MJO is much weaker and unrealistic compared to HadGEM1. However there are also some detriments: precipitation and temperature anomalies over Africa, S. America are made worse and the warm bias in summer over N. America and Asia is made worse. These appear to be related to issues with the current operational model and hence if the spring physics changes (model cycle G39) are introduced to the ensemble system we will be able to use future updates to model physics.

The main outstanding issues with the revised model version for 10-15 day forecasts are: large positive temperature biases in the tropical stratosphere; warm summer biases over Asia and North America; large positive biases in relative humidity near the tropopause; a strengthening of meridional circulation but also a confining of the errors to a narrower height range; a tendency for the 500 hPa geopotential heights to reflect a tendency to more zonal flow; an initial increase in available zonal potential energy and available eddy potential energy for winter cases; upper level convergence is too strong over the Pacific for summer; a positive bias in precipitation over Africa and northern parts of South America; a loss of predictability of blocking after 5 days with a tendency at 10-15 days for the North Atlantic blocking region to move westward; and too low an amplitude of the Madden-Julian Oscillation.

The size of the biases in the forecasts with the updated physics are in general of a similar magnitude to those of the control of the ECMWF ensemble prediction system. Many systematic errors show similar patterns between the two models. The most significant area in which the UM performs worse than ECMWF is a much larger summertime positive temperature bias over North America.

The large number of model changes made in this experiment makes it more difficult to find reasons for both improvements and detriments. It could prove useful to perform an additional set of experiments concentrating on the impact of specific model changes. In particular it would prove useful to investigate the impact of the soil moisture by rerunning the cases with the PS11 model configuration but using climatological soil moisture content fields. A package of changes aimed at reducing the summer warm bias based on changes made for HadGAM1a (including the convective cloud amount decay mentioned earlier) will be tested for 15 day forecasts soon.

References

Bowler, N., A. Arribas, K.R. Mylne and K.B. Robertson, (2007a), Met Office Global and Regional Ensemble Prediction System (MOGREPS) Part I: System description, Forecasting Research Technical Report in preparation.

Bowler, N., A. Arribas, K.R. Mylne, K.B. Robertson, S.E. John and Tim P. Legg, (2007b), Forecasting Research Technical Report in preparation.

Brown, A. R., and A. L. M. Grant, (1997), Non-local mixing of momentum the convective boundary layer. Boundary-Layer Meteorol., 84, 1-22.

Brown, A.R., Beare, R.J., Edwards, J.M., Lock, A.P., Keogh, S.J., Milton, S.F., and Walters, D.N. (2007), Upgrades to the boundary layer scheme in the Met Office NWP model. Submitted to Mon. Wea. Rev.

D'Andrea et. al., (1998), Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979-1988. Climate Dynamics, 14, 385-407.

Edwards, J. M., (2007), Oceanic latent heat fluxes: consistency with the atmospheric hydrological and energy cycles and GCM modelling. Accepted for publication in J. Geophys. Res.

Huffman, G.J., R.F. Adler, M.M. Morrissey, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, (2001). Global precipitation at one-degree daily resolution from multi-satellite observations. J. Hydrometeor., 2, 36-50.

Inness, P., 2007, The Madden-Julian Oscillation in HadGEM1a. Internal report.

James, P.M. (2006), An assessment of European synoptic variability in Hadley Centre Global Environmental models based on an objective classification of weather regimes, Climate Dynamics, 27 (2-3): 215-231.

Jones, C.P., (2004), Unified Model documentation paper no 70, Ancillary file data sources.

Jung T. and A. Tompkins, (2003), Systematic Errors in the ECMWF Forecasting System, Technical Memorandum 422, ECMWF.

Jung, T. (2003) Systematic errors of the atmospheric circulation in the ECMWF forecasting system, Q. J. R. Met. Soc., 131, 1045-1073.

Lau, K. and Waliser, D., (2005). Intra-seasonal variability in the atmosphere-ocean climate system. Springer-Praxis books in environmental science.

Li,D. and Shine, K.P. (1995). A 4-Dimensional Ozone Climatology for UGAMP Models. UGAMP Internal Report No. 35 (Unpublished paper).

Lorenc, A.C., S.P. Ballard, R.S. Bell, N.B. Ingleby, P.L.F. Andrews, D.M. Barker, J.R. Bray A.M. Clayton, T. Dalby, D. Li, T.J. Payne and F.W. Saunders (2000). The Met. Office global three-dimensional variational data assimilation scheme . Q. J. R. Met. Soc., 126 (570) 2991-3012.

Lorenz, E.N., (1967): The nature and theory of the general circulation of the atmosphere. Number 218.TP.115. World Meteorological Organisation, 159 pp.

Maidens A.V. and S.H. Derbyshire, (2007). Improving mass flux profiles in the Gregory-Rowntree Convection scheme using adaptive detrainment, manuscript in preparation.

Marshall, A., 2006, The Madden-Julian Oscillation in a modified version of the HADGEM1 model. Internal report.

Martin, G.M., M.A. Ringer, V.D. Pope, A. Jones, C. Dearden, and T.J. Hinton (2005). The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: Model description and global climatology, J. Climate, 19, 1274-1301.

Milton, S.F, Brooks, M., Lock, A., Whelan, E., Wilson,D., and Allan, R. (2005) HadGEM1 Physics for the Global NWP model (Cycle G34) : Improvements to Boundary Layer, Large Scale Precipitation, Convection and Saharan Albedo. Forecasting Research Technical Report 458.

Milton S.F., Hemming, J., Walters, D. and Willet, M. (2007). A preliminary assessment of the performance of the Global Unified Model in THORPEX 15-day forecasts, Forecasting Research Technical Report 495.

Parker, D.E, Jackson, M., Horton, E.B. (1995), The GISST2.2 sea surface temperature and sea ice climatology. Climate Research Technical Note 63. `http://www.metoffice.gov.uk/research/hadleycentre/obsdata/GISST.html`

Pelly, J.L. and B.J.Hoskins, (2003), A new perspective on blocking. J. Atmos. Sci , 60, 743-755.

Shapiro, M. and Thorpe, A. (2004). THORPEX: A global atmospheric research programme for the beginning of the 21st century, WMO Bulletin, July 2004 . `http://www.wmo.ch/thorpex/pdf/wmo_bul_fgge_July_2004.pdf`

Tibaldi, S. and F. Molteni, (1990), On the operational predictability of blocking. Tellus , 42A, 343-365.

Wheeler, M. and Hendon H. (2004). An all-season real time multivariate MJO index: development of an index for monitoring and prediction. Monthly Weather Review, 132, 1917-1932.

Wilson, M.F. and Henderson-Sellers, A. (1985). A global archive of land cover and soils data for use in general circulation climate models. Journal of Climatology, 5, pp119-143. `http://dss.ucar.edu/datasets/ds767.0/`

Willet, M, Unified Model has improved physics, NWP Gazette, June 2006, `http://www.metoffice.gov.uk/research/nwp/publications/nwp_gazette/jun06/um_imp_phys.html`.

Willmott, C.J., Rowe, C.M. and Mintz, Y. (1985). Climatology of the Terrestrial Seasonal Water Cycle, Journal of Climatology, 5, pp589-606.